

Db 462 CCGAGCGTCCGCTCAGAAAAGAGCTGGAAGTGGGCCCCAGCCCAAGCTGCCCCCTGGCC 521
QY 300 TGTGGCCCTGAGGCGCACTAGCTGCAGATGCTGAATGCTGTGGCCCACTGGCCCTCC 359
Db 522 TGTGGCCCTGAGGCGCACTAGCTGCAGATGCTGAATGCTGTGGCCCACTGGCCCTCC 581
QY 360 GTCTGGGCGCTTATGCTCTCTGAGGCGCGAGAGGCGGCGGCGCTTACAGAGCCCTGC 419
Db 582 GTCTGGGCGCTTATGCTCTCTGAGGCGCGAGAGGCGGCGGCGCTTACAGAGCCCTGC 641
QY 420 ACTGCGCTTACAGGCACTGAGTATPCTGCAAGGTGTACCCGCTGCAAGAAAGCCCTGGCC 479
Db 642 ACTGCGCTTACAGGCACTGAGTATPCTGCAAGGTGTACCCGCTGCAAGAAAGCCCTGGCC 701
QY 480 TGTGAGAGCCCTTACGAGGCGCTGCCCCGCAAGCAATGAGTGTGCTGGCGCCCACTGAGAGTCC 539
Db 702 TGTGAGAGCCCTTACGAGGCGCTGCCCCGCAAGCAATGAGTGTGCTGGCGCCCACTGAGAGTCC 761
QY 540 TGTGCTGTAACCAAGCTCTCTTCAAGCCCTTTTCACTGGAACCCATGGGGAATGCAAGCC 599
Db 762 TGTGCTGTAACCAAGCTCTCTTCAAGCCCTTTTCACTGGAACCCATGGGGAATGCAAGCC 821
QY 600 TGTGCTGTAACCAAGCTCTCTTCAAGCCCTTTTCAAGCCCTTTTCAAGCCCTTTTCAAGTGG 659
Db 822 TGTGCTGTAACCAAGCTCTCTTCAAGCCCTTTTCAAGCCCTTTTCAAGCCCTTTTCAAGTGG 881
QY 660 CCAAGCGCCCTGAGGCGCACTGTACAGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 719
Db 882 CCAAGCGCCCTGAGGCGCACTGTACAGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 941
QY 720 GCTTTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
Db 942 GCTTTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
QY 780 GCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839
Db 1002 GCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1061
QY 840 GACTGAGATATCTCAGCTCAAGGAGCTCATATCTCGGCGCAAGGCGAGCTGCTGAGAGCC 899
Db 1062 GACTGAGATATCTCAGCTCAAGGAGCTCATATCTCGGCGCAAGGCGAGCTGCTGAGAGCC 1121
QY 900 TGGGCGTGGGCGCTCTTCAAGGAGCTGAGGCGCACTAAGCCCTTCAAGGAGCTGAGAGCTG 959
Db 1122 TGGGCGTGGGCGCTCTTCAAGGAGCTGAGGCGCACTAAGCCCTTCAAGGAGCTGAGAGCTG 1181
QY 960 TCTGTCTTCTGCGCAAGTCCGCGCGGCGCTTACGCTTGTGAGAGCTTCTGAGAGCTTCTGAGCC 1019
Db 1182 TCTGTCTTCTGCGCAAGTCCGCGCGGCGCTTACGCTTGTGAGAGCTTCTGAGAGCTTCTGAGCC 1241
QY 1020 CTGCGCGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
Db 1242 CTGCGCGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1301
QY 1080 CAGGCAATCTCTCTGCAAGCCCTGCTGCGACAGAGCCGATGCTTACGCTTACGCTTACGCTTAC 1139
Db 1302 CAGGCAATCTCTCTGCAAGCCCTGCTGCGACAGAGCCGATGCTTACGCTTACGCTTACGCTTAC 1361
QY 1140 CCCATCTCTGCGAGGCTGCGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1199
Db 1362 CCCATCTCTGCGAGGCTGCGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1421
QY 1200 AAGAGGAGGAGAGCAGAGAGTGTCTGTATGCTGAGACCACTTACATACAGCTGAG 1259
Db 1422 AAGAGGAGGAGAGCAGAGAGTGTCTGTATGCTGAGACCACTTACATACAGCTGAG 1481
QY 1260 CTGCGCAACAGTGTGATTTGGGAGTGTGCTGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1319
Db 1482 CTGCGCAACAGTGTGATTTGGGAGTGTGCTGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1541
QY 1320 CCAAAAGCTTCAAGTGTGCTGCAAGAGGAGAAAGGCGAGAGCTGTGTGAGAGTGTGTG 1379
Db 1542 CCAAAAGCTTCAAGTGTGCTGCAAGAGGAGAAAGGCGAGAGCTGTGTGAGAGTGTGTG 1601

QY 1380 TACACATCTGCTTGTGTTCCACACATGCAAGTTCCTGCTTGGGCTGTATACAGGTGCCAA 1439
Db 1602 TACACATCTGCTTGTGTTCCACACATGCAAGTTCCTGCTTGGGCTGTATACAGGTGCCAA 1661
QY 1440 GCCCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1499
Db 1662 GCCCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1721
QY 1500 AGAGATGACAACTGCGATCTTGTAGCTGACAACTTCTTCAATGACATAGTCACTGT 1559
Db 1722 AGAGATGACAACTGCGATCTTGTAGCTGACAACTTCTTCAATGACATAGTCACTGT 1781
QY 1560 CTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1619
Db 1782 CTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1841
QY 1620 CTGTCCAAAGGACAACTGCTTTCACAAACCAAGCTGCTTGTATCTGTACCTTTTC 1679
Db 1842 CTGTCCAAAGGACAACTGCTTTCACAAACCAAGCTGCTTGTATCTGTACCTTTTC 1901
QY 1680 AGAGAAAGGAGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1739
Db 1902 AGAGAAAGGAGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1961
QY 1740 CAAAGCCAGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799
Db 1962 CAAAGCCAGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2021
QY 1800 TCTCTCTTCCAGGCTTAAAGCTTGGGATTTGGGCGCAGAGATGAGATCCAACTATGAGGCT 1859
Db 2022 TCTCTCTTCCAGGCTTAAAGCTTGGGATTTGGGCGCAGAGATGAGATCCAACTATGAGGCT 2081
QY 1860 AGTTCTTGTCTAATCTAAGCTGTCTTGTGAAATGAGGCTCAGGCTGTCAACATGAGGCT 1919
Db 2082 AGTTCTTGTCTAATCTAAGCTGTCTTGTGAAATGAGGCTCAGGCTGTCAACATGAGGCT 2141
QY 1920 TTCTGACTGAGACCAAGGTTGAGGAGACAGATTTAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1979
Db 2142 TTCTGACTGAGACCAAGGTTGAGGAGACAGATTTAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2201
QY 1980 TGGAAAGTCCAGAGTGGAGCTCTTCTGCGGAGACACTTGGGAGTCCCAATCCAGAGTCCATA 2039
Db 2202 TGGAAAGTCCAGAGTGGAGCTCTTCTGCGGAGACACTTGGGAGTCCCAATCCAGAGTCCATA 2261
QY 2040 CTCTAGCTTTTGTGATCCATGAGTATGTTATCTTACCTGTGCTTATTAAGAGATTTAT 2099
Db 2262 CTCTAGCTTTTGTGATCCATGAGTATGTTATCTTACCTGTGCTTATTAAGAGATTTAT 2321
QY 2100 GAAATTAATAAAAAAAAAA 2116
Db 2322 GAAATTAATAAAAAAAAAA 2338

RESULT 2
US-09-509-902A-3
; Sequence 3, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509, 902A
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1085
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-509-902A-3

Query Match 50.6%; Score 1069.8; DB 4; Length 1085;
Best Local Similarity 99.7%; Pred. No. 4,5e-267;
Matches 1082; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
QY 35 CGGGCAGGGCTGGAGCTGGGATCCCGAGCTGGGAGAGGCA-CGGGCGGGCCC 93
DB 1 CGGGCAGGGCTGGAGCTGGGATCCCGAGCTGGGAGAGGCA-CGGGCGGGCCC 60
QY 94 ACCTGTGCTGCTGGAGAGCTCTGAGCCCGGCGGGCGCGGGGCGCAAGAGC 153
DB 61 ACCTGTGCTGCTGGAGAGCTCTGAGCCCGGCGGGCGCGGGGCGCAAGAGC 120
QY 154 GGGCGAGATCGAGCCGCTCTGTGCTCTGTGGGCTTCCCTGTCCAGGAGAGCG 213
DB 121 GGGCGAGATCGAGCCGCTCTGTGCTCTGTGGGCTTCCCTGTCCAGGAGAGCG 180
QY 214 GTTGAAGTTGATGACACTTATGATACGAGCGTCCCGTCCAGAAAAGAGTGG 273
DB 181 GTTGAAGTTGATGACACTTATGATACGAGCGTCCCGTCCAGAAAAGAGTGG 240
QY 274 GCGCCAGCCAGAGCGCCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 333
DB 241 GCGCCAGCCAGAGCGCCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 334 TGCACCTGTGTGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 393
DB 301 TGCACCTGTGTGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 394 GGGCGGGCGGGCTTACAGGCGCTTACAGGCGCTTACAGGCGCTTACAGGCG 453
DB 361 GGGCGGGCGGGCTTACAGGCGCTTACAGGCGCTTACAGGCGCTTACAGGCG 420
QY 454 GTACCCCGCTCAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 513
DB 421 GTACCCCGCTCAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 514 GCATGTGCTCGGCGCTGAGAGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 573
DB 481 GCATGTGCTCGGCGCTGAGAGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 574 TTGGAACCATGGGAGATGACAGGCTGTGTGGAAGCGCGCACCGTATCCCTGA 633
DB 541 TTGGAACCATGGGAGATGACAGGCTGTGTGGAAGCGCGCACCGTATCCCTGA 600
QY 634 GGGTGCCTGCTCTTCCGCGAGATGGCCAGCGCGCGCGCGCGCTGCAAGAG 693
DB 601 GGGTGCCTGCTCTTCCGCGAGATGGCCAGCGCGCGCGCGCGCTGCAAGAG 660
QY 694 GGTCTGCTGCTGCTCAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 753
DB 661 GGTCTGCTGCTGCTCAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 754 GGTGCTGAGAACTTGGAGAGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 813
DB 721 GGTGCTGAGAACTTGGAGAGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 814 CAGAGACGCGTGCAGGCTTACGTTGAGACTGAGATCACTGAGTCAAGGCTCA 873
DB 781 CAGAGACGCGTGCAGGCTTACGTTGAGACTGAGATCACTGAGTCAAGGCTCA 840
QY 874 GGGCAGAGGAGCGGATGTGTGAGCGCTGGGCGGTGCGCTTCAAGCATCTGG 933
DB 841 GGGCAGAGGAGCGGATGTGTGAGCGCTGGGCGGTGCGCTTCAAGCATCTGG 900
QY 934 CTAACCTTTCAGAGACTCGGAGCTGTCTGTCTTCCGCAAGATCCGCGGGGCT 993
DB 901 CTAACCTTTCAGAGACTCGGAGCTGTCTGTCTTCCGCAAGATCCGCGGGGCT 960
QY 994 CGCTTGTGCTGAGAGGCTGTGCGGCGCTGCGGCGCTGCTGCTGCTGCTGCTG 1053
DB 961 CGCTTGTGCTGAGAGGCTGTGCGGCGCTGCGGCGCTGCTGCTGCTGCTGCTG 1020
```

```
QY 1054 GAGCGACCTGAAGCGCTGCAAGCCACAGGCAATCTCTGCAACCCCTGCTGCA 1113
DB 1021 GAGCGACCTGAAGCGCTGCAAGCCACAGGCAATCTCTCTGCAACCCCTGCTGCA 1080
QY 1114 CCGCA 1118
DB 1081 CCGCA 1085
```

RESULT 3

US-09-799-875-9

; Sequence 9, Application US/09799875

; Patent No. 6638721

; GENERAL INFORMATION:

; APPLICANT: Meyers, Rachel

; APPLICANT: Kapeller-Libermann, Rosana

; APPLICANT: Williamson, Mark

; TITLE OF INVENTION: No. 6638721e1 Human Protein Kinases and Uses

; FILE REFERENCE: 35800/209996

; CURRENT FILING DATE: US/09/799,875

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

; PRIOR FILING DATE: 2000-02-11

[illegible]

QY	836	GTGGGACCTTGAGATACACTACGCTCAACGGGCTCTCAACTCGGGCAAGGACGGCATGTCTGG	895
Db	1515	GTGGGACCTTGAGATACACTACGCTCAACGGGCTCTCAACTCGGGCAAGGACGGCATGTCTGG	1455
QY	896	AGCCTGGGCGTGGCGGCTCTTACCATGTGGCCGGACCTAACCCCTTCAGAGACTGGAG	955
Db	1455	AGCCTGGGCGGCGGCTCTTACCATGTGGCCGGACCTAACCCCTTCAGAGACTGGAG	1356
QY	956	CCTGTCCGCTCTTCGGCAAGATCCGGCGGGGGCTTAACGCTTGGCTCGACGGGCTCTCG	1015
Db	1395	CCTGTCTGCTCTTTCGGCAAGATCCGGCGGGGGCTTAACGCTTGGCTCGACGGGCTCTCG	1336
QY	1016	GCCCCGACCGGCTGTGTGTGGTTCGCTGCTCTTCCTGTGTGGGAGCCAGCTGAACGGCTCA	10797
Db	1335	CCCTCGCCGGGCTGTGTGTGTTCGGTGCCTCTTCGTGTGGGAGCCAGCTGAACGGCTCA	1276
QY	1076	GCCACAGGCACTCTCTCTGCAACCCCTGGCTGGACAGAACCCGATGCCCTTAAGCCCAAC	1135
Db	1275	G-CACACACGATCCCTCTGCAACCCCTGGCTGGACA-GACCAGATGCCCTTAAGCCCAAC	1218
QY	1136	CGATCCCATCTCTTGGGAGGCTGCCAGGTGTTCCTGATGGAATGAGCGGGCTGAAGAAC	1195
Db	1217	CGATCCCATCTCTTGGGAGGCTGCCAGGTGTTCCTGATGGAATGAGCGGGCTGAAGAAC	1155
QY	1196	AGGGAGAAGGAGGGAGACAGAGAAAGTGTTCGTATGAGCTAGAACCAACCTTACTACAGC	1255
Db	1158	AGGGAGAAGGAGGGAGACAGAGAAAGTGTTCGTATGAGCTAGAACCAACCTTACTACAGC	1099
QY	1256	TCACCTCCCAACAGTGAATTGATTTGGGGGTACTCCAGAGCCTTCCCTGSCCTGTA	1313
Db	1098	TCACCTCCCAACAGTGAATTGATTTGGGGGTACTCCAGAGCCTTCTCTGSCCTGTA	1041

[illegible]

RESULT 5
 US-08-146-421-4/c
 Sequence 4 Application US/08146421
 Patent No. 5543499
 GENERAL INFORMATION:
 APPLICANT: BREWER, GARY
 TITLE OF INVENTION: DNA SEQUENCE ENCODING A POLYPEPTIDE WITH
 TITLE OF INVENTION: ANTI-TUMOR PROPERTIES
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DILMORTH & BARRESE
 STREET: 4350 LA JOLLA VILLAGE DRIVE, SUITE 300
 CITY: SAN DIEGO
 STATE: CALIFORNIA
 COUNTRY: U.S.A.
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/146,421
 FILING DATE: 29-OCT-1993
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: PEPPER PH.D., FREDERICK W.
 REGISTRATION NUMBER: 31,286
 REFERENCE/DOCKET NUMBER: 489-2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619-546-4410
 TELEFAX: 619-453-2839
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2562 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA

8/6/96

CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1146:
SEQUENCE CHARACTERISTICS:
LENGTH: 1302 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1480860
US-09-016-434-1146

Query Match 4.0%; Score 85; DB 4; Length 1302;
Best Local Similarity 48.1%; Pred. No. 2, 5e-12;
Matches 241; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

QY 615 ACCGTATCCCTGAGCTGAGCGTGGCTGCTTTCCGCGATGCGCACCGCCCTGCGC 674
DB 437 AGCGTTCCAGTGTGCGACGCGCCACGGGTACTTGTGATGATGACGGCTGGAGT 496
QY 675 ACTGTACACGACGCGGTGCTGCTGCGATCTCAAGCTGTGCTTGTGCTGCTG 734
DB 497 ACCTGATAGCGCGGCGATTTGTCAACAAGACATCAACCGGGGAACTGCTCACC 556
QY 735 ACCGTAGAGGAAGAGCTGTGCTGAGAACCTGAGAGACTCTGCGGTGAGTGGGC 794
DB 557 CCGGTGGACCCCTCAAAATCTCCGACCTGGCGCTGGCCAGAGCATGCACTCCG 616
QY 795 CAGATATTCCTGTTGGGACAGACGCGTGGCCAGCTTACGTGGACCTGAGATCT 854
DB 617 CGACACACCTGCGCGACACGACGAGCGCTCCCGGCTTCCAGCGCCGAGATTG 676
QY 855 GCTACGCGGCTCATTACTCGGCGAAGGCAAGCGATGCTGGAGACCTGGCGGTG 914
DB 677 ACGGCTGACACCTTCTCCGCTTAAGGTGACATCTGGTGGGTGACCTCTCT 736
QY 915 TACCATGTGCGCGGCACTACCCCTTCAGAGACTCGAGCCTGCTCTTGGGCA 974
DB 737 ACAACATCAACAAGGCTGTGACCTCTTGAAGGGGACAAACATCTCAAGTTGTT 796
QY 975 AATTCGCGCGGCGCTTACGCTTTCCTGCAAGCTCTTGGCGCTGCGCTGTG 1034
DB 797 ACATCGGGAAGGAGGAGCTACGCGCATCCGCGGAGACTGTGGCCCGCTCTG 856
QY 1035 TTCGCTGCTCTTGTGGGAGGAGCAAGCGGTCAAGCAGACAGCATCTCTCTG 1094
DB 857 TGAAGGATGCTTGAATGAGAACCGGCGCAAGAGGTTCTTCATCGGGAGATCCG 916

QY 1095 ACCCTGCTCGACAGACC 1115
DB 917 ACAGCTGTTCCGAGAGAAC 937

RESULT 8

US-08-749-902-2
Sequence 2, Application US/08749902
Patent No. 5985635
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goll, Surya K.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL HUMAN SERINE/THREONINE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,902
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0150 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-4166
TELEFAX: 415-855-0555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1466 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: Consensus
US-08-749-902-2

Query Match 3.9%; Score 81.6; DB 2; Length 1466;
Best Local Similarity 47.3%; Pred. No. 2e-11;
Matches 237; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 615 ACCGTATCCCTGAGCTGAGCGTGGCTGCTTTCCGCGATGCGCACCGCCCTGCGC 674
DB 536 AGCGTTCCAGTGTGCGACGCGCCACGGGTACTTGTGATGATGACGGCTGGAGT 595
QY 675 ACTGTACACGACGCGTGTGCTGCGATCTCAAGCTGTGCTTGTGCTGCTG 734
DB 596 ACCTGATAGCGAGGATTTGTGACAGAGGATCAAGCCGGGAACTGCTGCTACCA 655
QY 735 ACCGTAGAGGAAGAGCTGTGCTGAGAACCTGAGAGACTCTGCGCTGACTGGGC 794
DB 656 CCGGTGGACCCCTCAAAATCTCCGACCTGGGCGTGGCCAGGACATGCACTCCG 715
QY 795 CAGATATTCCTGTTGGGACAGACGCGTGGCCAGCTTACGTGGACCTGAGATCT 854
DB 716 CGAGACACCTGCGGAGCAGCCGAGGCTCCCGGCTTTCAGCGCCCGGAGATTG 775
QY 855 GCTACGCGGCTCATTACTCGGCGAAGGCAAGCGATGCTGGAGCTGGCGGTG 914

Db 776 ACGGCTGACACCTTCTCCGCTTCAAGGTGACATCTGTCGCTGGGAGTACCTCT 835
Qy 915 TCACCATGCTGGCGCCGACACTACCCCTTCAGAGACTCGAGGCTGTCTCTTCGGA 974
Db 836 ABAACATCAACAGGGTCTGTACCTTTCGAGGGGACACAATCTACAGTTGTTGAA 895
Qy 975 AATTCGCGCGGGGCTTACGCTTGTGACAGGCTTCGAGCCCTGCGCCGCTGTG 1034
Db 896 ACATCGGGAAGGAGCTACGCTACGCTCCGCGACTGTGGCCCTCCGCTCTGACCTG 955
Qy 1035 TTGCTGCTCTCTTCTGCTGGAGCCAGCTGAAGGCTCACAGCCACAGCATCTCTG 1094
Db 956 TGAAGAGATGCTTGTAGTACGAACCGGCCAAGAGTTCTTCATCGGAGATCCGAGC 1015
Qy 1095 ACCCTGCTGCGACGACGACC 1115
Db 1016 ACAGCTGTTCGAGAGAAC 1036

RESULT 9
US-09-016-434-391
; Sequence 391, Application US/09016434
; Patent No. 6500938

GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 391:
SEQUENCE CHARACTERISTICS:
LENGTH: 288 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITW03
CLONE: 2108752
US-09-016-434-391

Query Match 3.4%; Score 71; DB 4; Length 288;
Best Local Similarity 63.3%; Pred. No. 5.6e-09;
Matches 126; Conservative 0; Mismatches 70; Indels 3; Gaps 1;
Qy 806 CTGTGGACAGACGGGTGCCCAAGCTTACGTGGAGCTGAGTACTACGCTACAGGGCC 865

Db 69 CTGAAGACCAAGAGGAGGAGCCCTTCTACATCACTCCGACGTGCTACGCGCGG--- 125
Qy 866 TCATACCTGGGGAAGGACAGCCGATGTCTGAGCCTGGGCGTGGCCCTTTCAACATGCTG 925
Db 126 CCGTACCGTGGGAGACCCAGTACATGTGGGCGCTGGGGGTGTCTCTTACACATGCTG 185
Qy 926 GCGGCACTACCCCTTCAGAGCTCGAGCCTGTCTCTTCTGGCAAGATCCGCGC 985
Db 186 TATGCGCATTCCTCTTACAGACGATCCGCGAGAGCTCTTCGCAAGATCAAGGCT 245
Qy 986 GGGGCTTACGCTTGTCTG 1004
Db 246 GCGGATATACCATTCCTG 264

RESULT 10
US-09-016-434-1147
; Sequence 1147, Application US/09016434
; Patent No. 6500938

GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1147:
SEQUENCE CHARACTERISTICS:
LENGTH: 2169 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 91488262
US-09-016-434-1147

Query Match 3.1%; Score 65.6; DB 4; Length 2169;
Best Local Similarity 45.9%; Pred. No. 3.2e-07;
Matches 310; Conservative 0; Mismatches 354; Indels 12; Gaps 2;
Qy 448 CAGGTGTACCCCGTCCAGAAACCCCTGCGCTGAGCCCTTACGCGGCTGCCCC 507
Db 216 CAAGCGCATCAAGCGGAGAGATCTTAATGATGATGAGCTGACCGAGACTTGAGCA 275
Qy 508 GCACAGCATGTGTGCTGCGCCCACTGAGGTCTGTGTGTATCCCAAGTCTCTACGCTT 567

Db 276 CCGGCATGCTGCTGTTTTCGACCACTTGAAGAGCGTGAACAACATTAATTTCTT 335
Qy 568 TTTCACCTGACCCCATGAGGAGATGACAGCTGTGTGGAACGGCCATCTCCTGA 627
Db 336 GAGAGCTTGAGCGGAAAGTCTCTGGCCCACTGTGAAGGCTCGGACACCCCTTTGA 395
Qy 628 GCGTGAAGCTGCGCTGCTCTTCCGCGAGATGGCCAGCCCTGGCGCACTGTCAACGCA 687
Db 396 GCGAGAGTGTGCTACTACCTGCGGAGATCTTCTTCTGCTCAAGTACTTGGACCGG 455
Qy 688 GCGTGTGTCTGCTGATCTCAAGCTGTGTCTGCTTGTCTTGTGCGTGAAGGAA 747
Db 456 GCGGATCTTGACCGGAGCTCAAGTGTGAATAATTTTCATCAC-----TGAGACAT 509
Qy 748 GAAGCTGTGTGAGAAACCTGAGGAGCTCCTGTGTCTGATGCGGCAATGATTCCT 807
Db 510 GGAATGAGAGTGGGGATTTTGGGCTGGCAAGCCGTTGAGCTCGGAGCAGAGAA 569
Qy 808 GTGGGACAGACAGCTGTGCGGAGCTGAGCTGAGACTGAGATACTAGCTCAGCGGCTC 867
Db 570 GAGAGCATGTGTGGACACCCCACTATGTGTCTCCAGAGTGTGTGAGACAGGCA 629
Qy 868 ATACTGCGGCAAGGAGCCGATGTCTGGAGCTGGGCGTGGCCCTTTCACATGTGCG 927
Db 630 C-----GGCCCTGAAGCGGATGTATGTCTAGTGGGCTGTGATGACAGCTGCTCTG 683
Qy 928 CCGGCATACACCCCTTCCAGAGCTCGGAGCTGTCTGTCTTGGCAAGATCCGCGCGG 987
Db 684 CCGGAGCTCTCCCTTTGAGAGCGCTGACGTGAAGAGACCTACCGCTGACAGAGGT 743
Qy 988 GGCCTAGCCTTGTCTGTGAGAGCTCTGCGCCCTGCGCTGTCTGTGTGCTGCTCT 1047
Db 744 TCACTACAGCGCTGCTGCGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 803
Qy 1048 TCGTGGGAGCGCTGAGAGGCTCACAGCGCAAGGATCTCTGACCCCTGCGTGG 1107
Db 804 TCGGAGCTGACCCCGAGAGCGCCCTCTATTGACAGATCTCGCATGCTTTTAC 863
Qy 1108 ACAGGACCGGATGCCC 1123
Db 864 CAGGCTACACCC 879

RESULT 11
US-09-579-664B-4
; Sequence 4, Application US/09579664B
; Patent No. 6514719
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MORINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/09/579,664B
; CURRENT FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2902
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-579-664B-4

Query Match 2.6%; Score 55.6; DB 4; Length 2902;
Best Local Similarity 49.0%; Pired. No. 0.00014;
Matches 244; Conservative 0; Mismatches 239; Indels 15; Gaps 3;

Qy 578 ACCCATGGGAGATGACAGCGCTGGCGAAGCGCCACCGATCTCTGAGCTGAGGCT 637
Db 531 AGCCGAGGAGATCTGTATATTATCATCATGAGAGCGGCCACGCTGAGTGGGAGCGCC 590

Qy 638 GCGGTCTCTTCGCGCAGATGAGCCAGCCCTGTGCGGACTGTACCAAGCAAGTGTGTC 697
Db 591 AGGATTTCTTCGACAGATGTGTCTGCTCTGACTACTACCAAGCAAGGATGTGT 650
Qy 698 CTGCGTATCTCAAGCTGTGTCTTGTCTTGTGCTGACCGTGAAGAGAAAGTGTG 757
Db 651 CACCGAGATCTCAAGCTGGAAGAAATCTCTTCTGATGACCAATGAAACATCAAGATTGCT 710
Qy 758 CTGAGAGCTGAGAGACTCTGCGTGTGCTGAGCTGGGCGAGATATTCCTGTGGACAG 817
Db 711 GACTTGGCT-----CTCAACCTGTACCAAGCAAGGAAAGTCTCTCAAGCGTTC 761
Qy 818 CAGCGGTCCAGCCTACGTTGAGGAGCTGTGATCTCACTGACGAGGCTTACTCGGC 877
Db 762 TGTGGAGCTCTCTTACGCTCTGCTGAGATATGTCAAGGAAAGCC---TATGTGGC 818
Qy 878 AAGCAGCCGATGTGAGGCTGTGGGCGTGGGCGCTCTTCACTGCTGGCGGCACTAC 937
Db 819 CAGAGGTGAGAGCTGTGTCTGCGGCTTCTCTGTACATCTGTGTGATGGACCATG 878
Qy 938 CCTTCCAGGACTGAGAGCTGTCTGCTTCTTGGCAAGATCCGCGGCGGCTTACGCC 997
Db 879 CCTTTGACCGGAGAGATCATAAACACTGTGTGAGCAATCAGTAACGGGCTTACCGT 938
Qy 998 TGGCTGAGAGGCTCTGCGGCGCTGCGGCGCTGTGCTGTGCTGCTCTGCTGCGGAG 1057
Db 939 GAGCGCC---CAAGCGTTCATGCTGCTGTGAGCTTATCGGTGCTTATGTGTAAC 995
Qy 1058 CAGCTGAACGCTCA 1075
Db 996 CCGACCGTGGGCGACA 1013

RESULT 12
US-08-125-468-1/c
; Sequence 1, Application US/08125468
; Patent No. 5589385
; GENERAL INFORMATION:
; APPLICANT: Ryan, Michael J.
; APPLICANT: Lotvin, Jason A.
; APPLICANT: Strathy, Nancy
; APPLICANT: Fancini, Susan E.
; TITLE OF INVENTION: Cloning of the biosynthetic pathway for
; TITLE OF INVENTION: chlorotetracycline and tetracycline Formation and cosmids
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07470
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,468
; FILING DATE: 22-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsevdos, Estelle J
; REGISTRATION NUMBER: 31,145
; REFERENCE/DOCKET NUMBER: 31,255-02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201)831-3241
; TELEFAX: (201)831-3305
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30001 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-125-468-1

Query Match 2.5%; Score 52.8; DB 1; Length 30001;
Best Local Similarity 44.4%; Pred. No. 0.0019;
Matches 213; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 550 CCAGCTCTCTACAGCCTTTTTCACCTGAGACCCATGGGGGACATGACAGCCTGTGGAGAC 609
DB 22561 CCAGAGGCTCAACGGGAGTTTGGCTTCCGCTCTGGGACGACACAGGAGAACTCTT 22502
QY 610 GGGCCACCGCTATCCCTGAGCCTGAGAGCTCCGCTGCTCTTCCGCAAGATGGCCACCGCCT 669
DB 22501 GCTGGTCCGCGACCGGATGGGCGTCAAGCCGCTCTACTACTGCGGACCGCCGCGCGCT 22442
QY 670 GGGCGACTGTCAACGACGAGGTCTGTGCTGCGGTATCTCAAGCTGTGTGCTTGTCTT 729
DB 22441 GGGCTTCCGCTCGAGACCCCAAGGCGCTCTGCGGCGACCGGACCTTCCCGCGGCTCGG 22382
QY 730 CGCTGACCTGTAGAGAGAAAGTGTGCTGTGAGAACTGTGAGAGACTCTGGTGTGTGAC 789
DB 22381 CCGGACGGGCTGTGTGAGAGTGTGACATGTGTGAAGACCGCCGAGGCGGCTCTTCTC 22322
QY 790 TGGGCGAGATGATTCCTGTGGGACAAAGACGCGTCCGACCTTACGTGGGAACTTGAGAT 849
DB 22321 CCGGCTCTGTGAGAGTCTGCGGATGAGCTCTGCGGATGGGCGGCGGCGGCTGGCGG 22262
QY 850 ACTGACCTGACGCGGCTCTACTACTGCGGCAAGGACCGGATGTGTGAGAGCTGTGGGCTGGC 909
DB 22261 GAGAGCTGACTGCGGCTGAGAGGCGCGGACGACACCGGACGACTGCGGACGACTCGG 22202
QY 910 GCTCTTACCATCTGTGCGGCGCACTACCTCTTCCAGAGACTCGGAGCTGTGCTGTCTT 969
DB 22201 CACCGTCCGAGACTGTGTGCGGACACCGTCACTGAGCTGTGCTGTGAGCTGTGCGCT 22142
QY 970 CGGCAAGATCCGCGCGGCGGCTTACGCGCTTGTGAGAGGCTCTCGGCGCTGCGCGCTG 1029
DB 22141 GGGCAGGCTGTCTCTCGGCGGCTGTGAGCTGTGCGGCTGTGAGCGGCTGTGCGCGCGG 22082

RESULT 13

US-08-474-933-1/C
Sequence 1, Application US/08474933
Patent No. 5866410
GENERAL INFORMATION:
APPLICANT: Ryan, Michael J.
APPLICANT: Lotvin, Jason A.
APPLICANT: Stralby, Nancy
APPLICANT: Parthini, Susan E.
TITLE OF INVENTION: Cloning of the biosynthetic pathway for
TITLE OF INVENTION: chlorotetracycline and tetracycline formation and cosmids
TITLE OF INVENTION: useful therein
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,933
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/125,468
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Tsevdos, Estelle J
REGISTRATION NUMBER: 31,145
REFERENCE/DOCKET NUMBER: 31,255-02
TELEPHONE: (201) 831-3241
TELEFAX: (201) 831-3305
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 30001 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-474-933-1

Query Match 2.5%; Score 52.8; DB 2; Length 30001;
Best Local Similarity 44.4%; Pred. No. 0.0019;
Matches 213; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 550 CCAGCTCTCTACAGCCTTTTTCACCTGAGACCCATGGGGGACATGACAGCCTGTGGAGAC 609
DB 22561 CCAGAGGCTCAACGGGAGTTTGGCTTCCGCTCTGGGACGACACAGGAGAACTCTT 22502
QY 610 GGGCCACCGCTATCCCTGAGCCTGAGAGCTCCGCTGCTCTTCCGCAAGATGGCCACCGCCT 669
DB 22501 GCTGGTCCGCGACCGGATGGGCGTCAAGCCGCTCTACTACTGCGGACCGCCGCGCGCT 22442
QY 670 GGGCGACTGTCAACGACGAGGTCTGTGCTGCGGTATCTCAAGCTGTGTGCTTGTCTT 729
DB 22441 GGGCTTCCGCTCGAGACCCCAAGGCGCTCTGCGGCGACCGGACCTTCCCGCGGCTCGG 22382
QY 730 CGCTGACCTGTAGAGAGAAAGTGTGCTGTGAGAACTGTGAGAGACTCTGTGCTGTGAC 789
DB 22381 CCGGACGGGCTGTGTGAGAGTGTGACATGTGTGAAGACCGCCGAGGCGGCTCTTCTC 22322
QY 790 TGGGCGAGATGATTCCTGTGGGACAAAGACGCGTCCGACCTTACGTGGGAACTTGAGAT 849
DB 22321 CCGGCTCTGTGAGAGTCTGCGGATGAGCTCTGCGGATGGGCGGCGGCGGCTGGCGG 22262
QY 850 ACTGACCTGACGCGGCTCTACTACTGCGGCAAGGACCGGATGTGTGAGAGCTGTGGGCTGGC 909
DB 22261 GAGAGCTGACTGCGGCTGAGAGGCGCGGACGACACCGGACGACTGCGGACGACTCGG 22202
QY 910 GCTCTTACCATCTGTGCGGCGCACTACCTCTTCCAGAGACTCGGAGCTGTGCTGTCTT 969
DB 22201 CACCGTCCGAGACTGTGTGCGGACACCGTCACTGAGCTGTGCTGTGAGCTGTGCGCT 22142
QY 970 CGGCAAGATCCGCGCGGCGGCTTACGCGCTTGTGAGAGGCTCTCGGCGCTGCGCGCTG 1029
DB 22141 GGGCAGGCTGTCTCTCGGCGGCTGTGAGCTGTGCGGCTGTGAGCGGCTGTGCGCGCGG 22082

RESULT 14

US-09-691-861A-14
Sequence 14, Application US/09691861A
Patent No. 6482935
GENERAL INFORMATION:
APPLICANT: Wei, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: C1000892
CURRENT APPLICATION NUMBER: US/09/691,861A
CURRENT FILING DATE: 2000-10-18
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 601
TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(601)
; OTHER INFORMATION: n = A,T,C or G
US-09-691-861A-14

Query Match
Best Local Similarity 49.7%; Score 51; DB 4; Length 601;
Matches 188; Conservative 0; Mismatches 181; Indels 9; Gaps 2;

QY 580 CCATGGGACATGACAGCCTGTGGGAAACGGCCACCGTATCCCTGAGCCTGAGGCTGC 639
DB 73 CCAAGGCGACCTCTCGAGTTAATCAAAACCGGGAGCCCTGCAATGAGACGAAAGCTCG 132
QY 640 CGTGCTCTTCGCGCATGAGCCACCGCCCTGAGCAGTGTCAACGACGAGCTGTGCT 699
DB 133 CAAGAAGTTCCACGAGCTTCTTGCCCATGAGTACCTGACACGACCTGAGAGCTGTCCA 192
QY 700 GCGTATCTCAAGCTGTGTGCTTTGCTTGTGCTGACCGTGAAGAAAGAGTGTGCT 759
DB 193 CCGGAGCTCAAGGTGTGACCAACCTTCTGACAGAGACTTCAACATCAAGCTGTCCGA 252
QY 760 GGAAGACCTGAGAGACTCTCGTGTGACTGGGCGCAAGATATCCCTGTGGGACAAAGCA 819
DB 253 CTTCAGCTTCTCCAAAGGCTGCTGCGGATGACAGTGTGAAATGCAATTAAGCAAGAC 312
QY 820 C-----GCGTCCAGCCTAAGTGGGACCTGAGACTCAGCTACAGGCGCTCATCTC 873
DB 313 CTTCGTGTGTGACACAGCGTATGCGGCCCGAGAGTGT---GCAAGGCAATTCCTTACA 369
QY 874 GGGCAAGGACCGCATGTGTGAGAGCTGGGCTGTGCGCTCTTCAACATGCTGCGCGCA 933
DB 370 GCCCAAGGTGTGACAGACTGTGAGACCTGAGCGTATCTCATCATGTGTGCGGCTC 429
QY 934 CTACCCCTTCAGGACTC 951
DB 430 CATGCCCTACGACGACTC 447

RESULT 15
US-09-691-861A-1
; Sequence 1, Application US/09691861A
; Patent No. 6482935
; GENERAL INFORMATION:
; APPLICANT: Wei, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000892
; CURRENT APPLICATION NUMBER: US/09/691,861A
; CURRENT FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-691-861A-1

Query Match
Best Local Similarity 49.7%; Score 50.4; DB 4; Length 1104;
Matches 188; Conservative 0; Mismatches 181; Indels 9; Gaps 2;

QY 580 CCATGGGACATGACAGCCTGTGGGAAACGGCCACCGTATCCCTGAGCCTGAGGCTGC 639
DB 282 CCAAGGCGACCTCTCGAGTTAATCAAAACCGGGAGCCCTGCAATGAGACGAAAGCTCG 341
QY 640 CGTGCTCTTCGCGCATGAGCCACCGCCCTGAGCAGTGTCAACGACGAGTGTGCT 699
DB 342 CAAGAAGTTCCACGAGCTTCTTGCCCATGAGTACCTGACACGACCTGAGAGCTGTCCA 401
QY 700 GCGTATCTCAAGCTGTGTGCTTTGCTTGTGCTGACCGTGAAGAAAGTGTGCT 759

DB 402 CCGGAGCTCAAGTGTGACCAACCTTCTTGACAGGACTTCAACATCAAGCTGTCCGA 461
QY 760 GGAAGACCTGAGAGACTCTTGTGCTGTGACTGTGGCCAGATATTCCTGTGGGACAAAGCA 819
DB 462 CTTCAGCTTCTCCAAAGCCTGCTGCGGATGACAGTGTGAAATGCAATTAAGCAAGAC 521
QY 820 C-----GCGTCCAGCCTAAGTGGGACCTGAGATTAATCAAGCTTCAAGGCTCATACTC 873
DB 522 CTTCGTGTGTGACACCAAGGTATGCGGCCCGAGAGTGT---GCAAGGCAATTCCTTACA 578
QY 874 GGGCAAGGACCGCATGTGTGAGAGCTGTGCGCTCTTCAACATGCTGAGCGCGCA 933
DB 579 GCCCAAGGTGTGACAGACTGTGAGCCTAAGGCTATCTCATCATGTGTGCGGCTC 638
QY 934 CTACCCCTTCAGGACTC 951
DB 639 CATGCCCTACGACGACTC 656

Search completed: August 29, 2004, 12:06:29
Job time : 159 secs

[illegible]

Db	282	TCCGAGACTGGAGCAGACGCCAGGGGCGCGCCACCTGCTGGTGGCTTGGAGGCTCTGGA	341
QY	120	GCCCCCGGGGCGCCCGGGCCCACGGGAACGACCGGGGCGAGATGCGAGCCACCTCTGCG	179
Db	342	GCCCCCGGGGCGCCCGGGCCCACGCGGAACGACCGGGGCGAGATGCGAGCCACCTCTGCG	401
QY	180	CTGCTCTGGGGGTTCCCTGTGTCAGAGAAAGGGTTGGAGTTGGATGAGAACCTTAATA	239
Db	402	CTGCTCTGGGGGTTCCCTGTGTCAGAGAAAGGGTTGGAGTTGGATGAGAACCTTAATA	461
QY	240	CCGAGGCTCCCGCTTCAGAAACGAGCTTGAAATGGGCCCAAGCCAGACTGCCCTGGCC	299
Db	462	CCGAGGCTCCCGCTTCAGAAACGAGCTTGAAATGGGCCCCAGCCAGACTGCCCTGGCC	521
QY	300	TGTTGCCCCCGAGCCCACTACAGCTCAGATGCTGCAATGCTGCAATGCTGTGAGCCACTGCTCCC	359
Db	522	TGTTGCCCCCGAGCCCACTACAGCTCAGATGCTGCAATGCTGCAATGCTGTGAGCCACTGCTCCC	581
QY	360	GTCTTGGGCCCTATATGCTCTCTGGAGCCCGAGAGGGGCGGGGCGGAGCTTACAGAGCCCTGCG	419
Db	582	GTCTTGGGCCCTATATGCTCTCTGGAGCCCGAGAGGGGCGGGGCGGAGCTTACAGAGCCCTGCG	641
QY	420	ACTGAGCCCTACAGGCACTGAGTATACCTGCAAGGTATACCCCTCCAGAAAGCCCTGGCGG	479
Db	642	ACTGAGCCCTACAGGCACTGAGTATACCTGCAAGGTATACCCCTCCAGAAAGCCCTGGCGG	701
QY	480	TGCTGAGGCCCTACAGCGCGGCTGCGCCCGCACAAGAGATGTGGCTGGGCCCACTGAGGCTGC	539
Db	702	TGCTGAGGCCCTACAGCGCGGCTGCGCCCGCACAAGAGATGTGGCTGGGCCCACTGAGGCTGC	761
QY	540	TGGCTGTATCCCAAGCTCTCTACGCTTTTTCATCTGGACCCATGAGGAACATGCAAGCC	599
Db	762	TGGCTGTATCCCAAGCTCTCTACGCTTTTTCATCTGGACCCATGAGGAACATGCAAGCC	821
QY	600	TGGTGCAGAACGCGCCACGCTATCCCTGAGCCCTGAGGCTGCGGTGCTCTTCCGCGCAGATGG	659
Db	822	TGGTGCAGAACGCGCCACGCTATCCCTGAGCCCTGAGGCTGCGGTGCTCTTCCGCGCAGATGG	881
QY	660	CCACGCGCTTGGGCGCATGTCAACGACGAGTCTGTGCTCTGCGTATCTCAAGCTGTGTC	719
Db	882	CCACGCGCTTGGGCGCATGTCAACGACGAGTCTGTGCTCTGCGTATCTCAAGCTGTGTC	941
QY	720	GCTTGTCTTGTGCTACGCTGAGAGAGAAAGCTGTGCTGTGGAAACTGAGAGCTCTCT	779
Db	942	GCTTGTCTTGTGCTACGCTGAGAGAGAAAGCTGTGCTGTGGAAACTGAGAGCTCTCT	1001
QY	780	GCGTGTGACTTGGGCGAGATGATTCCTGTGTGGACAAACAGCGCGCCAGGCTTACGTGGG	839
Db	1002	GCGTGTGACTTGGGCGAGATGATTCCTGTGTGGACAAACAGCGCGCCAGGCTTACGTGGG	1061
QY	840	GACCTGAGATCTACGCTCAAGGCGCTTATATCTGGGCGAAGCGAGCGATGTGGAAGCC	899
Db	1062	GACCTGAGATCTACGCTCAAGGCGCTTATATCTGGGCGAAGCGAGCGATGTGGAAGCC	1121
QY	900	TGGGCGTGGCGCTCTTCAACATGTCTGGGCGGCGCACTACCCCTTCAGGACTTGGAGCGCTG	959
Db	1122	TGGGCGTGGCGCTCTTCAACATGTGTGGGCGGCGCACTACCCCTTCAGGACTTGGAGCGCTG	1181
QY	960	TTCGTGCTTGGGAGAAATCCGCGCGGGGCGCTAAGCCTTGCCTGCAAGGCTCTGAGGCC	1019
Db	1182	TTCGTGCTTGGGAGAAATCCGCGCGGGGCGCTAAGCCTTGCCTGCAAGGCTCTGAGGCC	1241
QY	1020	CTGCGCGCTGTCTGTTGCTGTCGCTCCCTGTGCGGGAGCGAGCTGAACGGTTCACAGCA	1079
Db	1242	CTGCGCGCTGTCTGTTGCTGTCGCTCCCTGTGCGGGAGCGAGCTGAACGGTTCACAGCA	1301
QY	1080	CAGGCATCTTCTGCACCCCTTGGCTGGACAGAACCCGATGCCCTTACGCCCAACCCGAT	1139
Db	1302	CAGGCATCTTCTGCACCCCTTGGCTGGACAGAACCCGATGCCCTTACGCCCAACCCGAT	1361
QY	1140	CCCATCTTGTGGAGGCTGCGCCAGGTGGTCCCTATGAGACTGGGGCTGAGAGAACCGAGG	1199

Db	136	CCCATCTCTGGAGAGGCTGCCCAAGTGTGTCCTCGATGGACTGCGGGCTGAGCAAGCAAGG	1421
QY	1200	AAGAAGAGGAGACAGAGAAGTGGTTCTGTATGTGGTAGGACCCTTACTACAGCTCAG	1255
Db	1422	AAGAAGAGGAGAGACAGAGAAGTGGTTCTGTATGTGGTAGGACCCTTACTACAGCTCAG	1481
QY	1260	CTGCACCAAGTGGATTGAGTTTGGGGGTATGCTCCAAAGCTTCTCTGCTCTGAACTGAG	1319
Db	1482	CTGCACCAAGTGGATTGAGTTTGGGGGTATGCTCCAAAGCTTCTCTGCTCTGAACTGAG	1541
QY	1320	CCAAACCTTGAGTGCCTTCAGAAAGGAGAAAGGACAGAAAGCTGTGTGAGAGTGTCTG	1375
Db	1542	CCAAACCTTGAGTGCCTTCAGAAAGGAGAAAGGACAGAAAGCTGTGTGAGAGTGTCTG	1601
QY	1380	TACACATCTGCTTTGTTTCCACACACATGCAATGCTCTGCTGGGTCTTATCAGGTGCCAA	1438
Db	1602	TACACATCTGCTTTGTTTCCACACACATGCAATGCTCTGCTGGGTCTTATCAGGTGCCAA	1661
QY	1440	GCCCTGTCTCGGTGCTGGAGTACAGACAGTAGACCAAGAGACAAATATTCCTGCTCAC	1499
Db	1662	GCCCTGTCTCGGTGCTGGAGTACAGACAGTAGACCAAGAGACAAATATTCCTGCTCAC	1721
QY	1500	AAGAAGCAAACTGGCATCTTTAGCTGACAAACATTTTCCATGACCAATAGTCACTGT	1559
Db	1722	AAGAAGCAAACTGGCATCTTTAGCTGACAAACATTTTCCATGACCAATAGTCACTGT	1781
QY	1560	CTACACTGGGTACACTTTTGACACAGTGTCCGCTCCACTATAGTGGGGCTCAGGACCT	1619
Db	1782	CTACACTGGGTACACTTTTGACACAGTGTCCGCTCCACTATAGTGGGGCTCAGGACCT	1841
QY	1620	CTGTCCAGAGCAATCCCTTTCCAAACAAACAGTGCCTTTGTATCTTGATCCTTTTC	1679
Db	1842	CTGTCCAGAGCAATCCCTTTCCAAACAAACAGTGCCTTTGTATCTTGATCCTTTTC	1901
QY	1680	AAGAAGAGGAGAGTATCCCTGTGCAAAAGGCTCAGAGCTTCCCTCGCACTCAGGACC	1739
Db	1902	AAGAAGAGGAGAGTATCCCTGTGCAAAAGGCTCAGAGCTTCCCTCGCACTCAGGACC	1861
QY	1740	CAAGCCAGCTCATCTGGGAACTGTGTCCAGCATCTGTCTCTCTGATTAAGAGAT	1799
Db	1962	CAAGCCAGCTCATCTGGGAACTGTGTCCAGCATCTGTCTCTCTGATTAAGAGAT	2021
QY	1800	TCTCCTTCAGGCTTAAGCCCTGGGATTTGGGCCACAGATTAAGATTCACATAGAGCT	1859
Db	2022	TCTCCTTCAGGCTTAAGCCCTGGGATTTGGGCCACAGATTAAGATTCACATAGAGCT	2081
QY	1860	AGTCTGTGTCTAATCTCAAGACTGTTCTGGAATGAGGCTCCAGGCTGTCAACCATGGGCG	1919
Db	2082	AGTCTGTGTCTAATCTCAAGACTGTTCTGGAATGAGGCTCCAGGCTGTCAACCATGGGCG	2141
QY	1920	TTCTGACCTGACACCAAGGTTGAGGACAGGATTAAGCAGGAGTCTGTCTGTGGCCACC	1979
Db	2142	TTCTGACCTGACACCAAGGTTGAGGACAGGATTAAGCAGGAGTCTGTCTGTGGCCACC	2201
QY	1980	TGGAAGTCCAGAGTGGGACTCTTCTGGGGACACTTGGGGTCCACAAATCCAGAGTCCATA	2039
Db	2202	TGGAAGTCCAGAGTGGGACTCTTCTGGGGACACTTGGGGTCCACAAATCCAGAGTCCATA	2261
QY	2040	CTCTAGGTTTTCATTCATCAGATGATGATGTTTACCTGTGCTAATTAAGAGATTAAT	2099
Db	2262	CTCTAGGTTTTCATTCATCAGATGATGATGTTTACCTGTGCTAATTAAGAGATTAAT	2321
QY	2100	GAATTAATAAAAAAAAAA 2116	
Db	2322	GAATTAATAAAAAAAAAA 2338	
RESULT 2			
US-10-649-156-7			
; Sequence 7, Application US/10649156			
; Publication No. US2004008346A1			
; GENERAL INFORMATION:			
; APPLICANT: Meyers, Rachel			


```
/ APPLICANT: Kapeller-Libermann, Rosana
/ APPLICANT: Williamson, Mark
/ TITLE OF INVENTION: No. US20040038346A1el Human Protein Kinases and Uses
/ TITLE OF INVENTION: Therefor
/ FILE REFERENCE: 35800/209996
/ CURRENT APPLICATION NUMBER: US/10/649,156
/ CURRENT FILING DATE: 2003-08-27
/ PRIOR APPLICATION NUMBER: US/09/799,875
/ PRIOR FILING DATE: 2001-03-06
/ PRIOR APPLICATION NUMBER: 60/182,059
/ PRIOR FILING DATE: 2000-02-11
/ PRIOR APPLICATION NUMBER: 09/659,287
/ PRIOR FILING DATE: 2000-09-12
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7
/ LENGTH: 2389
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (383)..(1456)
/ US-10-649-156-7

Query Match      99.3%; Score 2100.2; DB 13; Length 2389;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2113; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY      1  GGAGGCGGCTCCGCGGCGCTCCGCTGTGAGACCCGCGGAGGCGTGGAGCTGGGCTGGGA 60
DB      222  GGAGGCGGCTCCGCGGCGCTCCGCTGTGAGACCCGCGGAGGCGTGGAGCTGGGCTGGGA 281
QY      61  TCCCGAGCTCCGCGAGAGGCA-CGGGCGGCGCCACCTGCTGGTCCCTGAGGCTCTGA 119
DB      282  TCCCGAGCTCCGCGAGAGGCA-CGGGCGGCGCCACCTGCTGGTCCCTGAGGCTCTGA 341
QY      120  GCCCGGCGGCGCGCGGCGCCAGCGGAAAGCGGCGGAGATGCAAGCCACCTCTGTG 179
DB      342  GCCCGGCGGCGCGCGGCGCCAGCGGAAAGCGGCGGAGATGCAAGCCACCTCTGTG 401
QY      180  CTGCTCTCTGGGCTTCCCTGTCCAGAGAAAGCGGTTGGATGGATGCAACTAGAT 239
DB      402  CTGCTCTCTGGGCTTCCCTGTCCAGAGAAAGCGGTTGGATGGATGCAACTAGAT 461
QY      240  CCGAGGCTCCGCTCCAGAAAGAGCTGGAAGTGGGCCCAAGCCAGCTGCCCTGCC 299
DB      462  CCGAGGCTCCGCTCCAGAAAGAGCTGGAAGTGGGCCCAAGCCAGCTGCCCTGCC 521
QY      300  TGTGACCTTGAGCCCACTATGCTCCAGATGTGCACTGCTGTGGCACTGTCTCC 359
DB      522  TGTGACCTTGAGCCCACTATGCTCCAGATGTGCACTGCTGTGGCACTGTCTCC 581
QY      360  GTCTTGGGCTTATGTCTCTCTGAGGCCGAGAGGCGCGGCGCTTACAGGCGCTGC 419
DB      582  GTCTTGGGCTTATGTCTCTCTGAGGCCGAGAGGCGCGGCGCTTACAGGCGCTGC 641
QY      420  ACTGCGCTTACAGGCAAGATATCTGCAAGGTGTAACCCGATCCAGAAAGCCCTG 479
DB      642  ACTGCGCTTACAGGCAAGATATCTGCAAGGTGTAACCCGATCCAGAAAGCCCTG 701
QY      480  TGCTGAGAGCTTACGCGCGGCTGCCCCGCAAGAGATGTGCTGGGCCCACTGAG 539
DB      702  TGCTGAGAGCTTACGCGCGGCTGCCCCGCAAGAGATGTGCTGGGCCCACTGAG 761
QY      540  TGGCTGATCCAGCTCTCTTAAGCTTTTCACTCCGACCCATGGGAGCATGCAAGCC 599
DB      762  TGGCTGATCCAGCTCTCTTAAGCTTTTCACTCCGACCCATGGGAGCATGCAAGCC 821
QY      600  TGGTGCAAGCGGCGAGCTTCCCTGAGCTGAGGCTGCGGTGCTTTCCGCAAGATG 659
DB      822  TGGTGCAAGCGGCGAGCTTCCCTGAGCTGAGGCTGCGGTGCTTTCCGCAAGATG 881
QY      660  CCACCGCCTGGCGCACTGTCAACAAGAGGCTGGCTGTGCTGTGATCTCAAGCTGT 719
```

```
DB      882  CCACCGCCTGGCGCACTGTCAACAAGAGGCTGGCTGTGCTGTGATCTCAAGCTGT 941
QY      720  GCTTGTCTTCTGACCGCTGAGAGAAAGCTGTGCTGAGAACTCTGAGAGCTCT 779
DB      942  GCTTGTCTTCTGACCGCTGAGAGAAAGCTGTGCTGAGAACTCTGAGAGCTCT 1001
QY      780  GCGTGTCTGAGGCGCAGATATTCCTGTGGGCAAGCAGCGGCGCAGCTGCTG 839
DB      1002  GCGTGTCTGAGGCGCAGATATTCCTGTGGGCAAGCAGCGGCGCAGCTGCTG 1061
QY      840  GACCTGAGATCTAGCTCAGCGGCTCATCTCGGCGAGGCAAGCTGATGTTCAG 899
DB      1062  GACCTGAGATCTAGCTCAGCGGCTCATCTCGGCGAGGCAAGCTGATGTTCAG 1121
QY      900  TGGGCGGCGCTTTCACCATGCTGCGCGCACTACCTCTTCCAGGACTCGGCG 959
DB      1122  TGGGCGGCGCTTTCACCATGCTGCGCGCACTACCTCTTCCAGGACTCGGCG 1181
QY      960  TCTGTCTCTTCCGCAAGATCCGCGGCGGCGCTTACGCTTGGCTGAGGCTCTG 1019
DB      1182  TCTGTCTCTTCCGCAAGATCCGCGGCGGCGCTTACGCTTGGCTGAGGCTCTG 1241
QY      1020  CTGCGCGCTGTGCTGCTGCTGCTCTCTTCTGCGGAGCGCAGCTGAACGCTCA 1079
DB      1242  CTGCGCGCTGTGCTGCTGCTGCTCTCTTCTGCGGAGCGCAGCTGAACGCTCA 1301
QY      1080  CAGGCACTCTCTGCAACCTCTGCTGCGCAAGAGACCCATGCGCTTACGCCAG 1139
DB      1302  CAGGCACTCTCTGCAACCTCTGCTGCGCAAGAGACCCATGCGCTTACGCCAG 1361
QY      1140  CCCATCTCTGAGAGGCTGCGCGCAGGCTGCTGATGGACTGGGCGCTGAGCAAG 1199
DB      1362  CCCATCTCTGAGAGGCTGCGCGCAGGCTGCTGATGGACTGGGCGCTGAGCAAG 1421
QY      1200  AAGAGAGGAGAGACAGAGAGTGTCTGTATGTGATAGACCACTTACTACGCTCA 1259
DB      1422  AAGAGAGGAGAGACAGAGAGTGTCTGTATGTGATAGACCACTTACTACGCTCA 1481
QY      1260  CTGCAACAGATGATGATTTGGGGGTGACTCAAGCTTCTCTGCTGCTTCAACTG 1319
DB      1482  CTGCAACAGATGATGATTTGGGGGTGACTCAAGCTTCTCTGCTGCTTCAACTG 1541
QY      1320  CCAAACCTTCAAGCTTCCAGAGAGGAGAAAGGCAAGCTGTGAGTGTGCTG 1379
DB      1542  CCAAACCTTCAAGCTTCCAGAGAGGAGAAAGGCAAGCTGTGAGTGTGCTG 1601
QY      1380  TACACATCTGCTTGTTCACACATGCACTGCTGTGGTGTATCAGGTGCCAA 1439
DB      1602  TACACATCTGCTTGTTCACACATGCACTGCTGTGGTGTATCAGGTGCCAA 1661
QY      1440  GCCCTGTTCTGGTCTGGAGATACAGCAGTGAAGAAAGACATATTCCTGCTCAC 1499
DB      1662  GCCCTGTTCTGGTCTGGAGATACAGCAGTGAAGAAAGACATATTCCTGCTCAC 1721
QY      1500  AAGAGATGAACATCGATCTTGAAGCTGACCAACTTTCCTCAATAGCTCACTGT 1559
DB      1722  AAGAGATGAACATCGATCTTGAAGCTGACCAACTTTCCTCAATAGCTCACTGT 1781
QY      1560  CTACACTGGGTACATTTGTACAGAGTGGGCTTCACTGATGTGTCTCAGGACCT 1619
DB      1782  CTACACTGGGTACATTTGTACAGAGTGGGCTTCACTGATGTGTCTCAGGACCT 1841
QY      1620  CTGTCAAGAGCAATCCCTTTCACAAACAAACAGCTGCTTGTATCTGTACTTTC 1679
DB      1842  CTGTCAAGAGCAATCCCTTTCACAAACAAACAGCTGCTTGTATCTGTACTTTC 1901
QY      1680  AGAAGAAAGGAGATATCCGTGCAAGAGCTCCAGGCTCTCCCTGCAACTCAGAG 1739
DB      1902  AGAAGAAAGGAGATATCCGTGCAAGAGCTCCAGGCTCTCCCTGCAACTCAGAG 1961
QY      1740  CAAAGCAAGCTCACTGTGGAACTGTGTCCAGCATCTGTCTCTTGTATTAAGAT 1799
```


Db 1321 AGGGAAGAAAGCAGAGAGCCTGTGTGTGAGTGTGTGTGATACACATCTGCTTTGTTCCACAC 1380
QY 1403 ACATGCACTTCTGCTTGGGTGTGCTTATCAGGTGCCAAGCCCTGTTCTCGTGTCTGGAGT 1462
Db 1381 ACATGCACTTCTGCTTGGGTGTGCTTATCAGGTGCCAAGCCCTGTTCTCGTGTCTGGAGT 1440
QY 1463 ACAGAGTGTGAGAAAGAGACATATTCCTGCTCTCAGAGATGACAACTGTGGCATCCTT 1522
Db 1441 ACAGAGTGTGAGAAAGAGACATATTCCTGCTCTCAGAGATGACAACTGTGGCATCCTT 1500
QY 1523 GAGCTGACAAACATTTTCCATGACCAATAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1582
Db 1501 GAGCTGACAAACATTTTCCATGACCAATAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
QY 1583 AGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1642
Db 1561 AGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
QY 1643 CAAACAAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1702
Db 1621 CAAACAAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
QY 1703 CCAAAAGCTCCAGGCTCTCTCCCTGCACTGAGGACCCAAAGCCCAAGCTCACTGTGGAAAC 1762
Db 1681 CCAAAAGCTCCAGGCTCTCTCCCTGCACTGAGGACCCAAAGCCCAAGCTCACTGTGGAAAC 1740
QY 1763 TGTGTCCCAAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1822
Db 1741 TGTGTCCCAAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
QY 1823 GATTGTGGCCAGAGATTAAGATCAAACTATGAGGCTAGTCTTGTCTTAATCTCAAGACTG 1882
Db 1801 GATTGTGGCCAGAGATTAAGATCAAACTATGAGGCTAGTCTTGTCTTAATCTCAAGACTG 1860
QY 1883 TTCTGGAATGAGGCTTCCAGGCTGTCTGCACTGAGGCTTGTGACTGAGGACCAAGGTTG 1942
Db 1861 TTCTGGAATGAGGCTTCCAGGCTGTCTGCACTGAGGCTTGTGACTGAGGACCAAGGTTG 1920
QY 1943 AGGGAACAGATTAAGGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2002
Db 1921 AGGGAACAGATTAAGGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
QY 2003 TCTGGGACACTTGGGGTCCACAAATCCAGGCTCAATCTGAGTTTGTGATACCATGAG 2062
Db 1981 TCTGGGACACTTGGGGTCCACAAATCCAGGCTCAATCTGAGTTTGTGATACCATGAG 2040
QY 2063 TATGTATGTTTACCTGTGCTTATTAAGAGATTTATGAAAT 2104
Db 2041 TATGTATGTTTACCTGTGCTTATTAAGAGATTTATGAAAT 2082

RESULT 4
US-10-425-114-18645
; Sequence 18645, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 18645
; LENGTH: 2070
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

OTHER INFORMATION: Clone ID: LIB3101-194-B6_FLI
US-10-425-114-18645
Query Match 96.7%; Score 2046; DB 13; Length 2070;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2060; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 42 GAGTGTGAGCTGTGGGCTGTGGATATCCAGGCTGCGGACAGAGGCA-CGGGCGGGCCACTGTGT 100
Db 1 GAGTGTGAGCTGTGGGCTGTGGATATCCAGGCTGCGGACAGAGGCA-CGGGCGGGCCACTGTGT 60
QY 101 GGTGTGCTGAGAGGCTGTGAGGCTGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 160
Db 61 GGTGTGCTGAGAGGCTGTGAGGCTGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 161 ATGCGAGCACCCCTCTGAGCTGCTCTGCGGAGTTCCTGTGTCAGGAAAGAGCGGTTGAG 220
Db 121 ATGCGAGCACCCCTCTGAGCTGCTCTGCGGAGTTCCTGTGTCAGGAAAGAGCGGTTGAG 180
QY 221 TTGATGACACTTATGATACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 280
Db 181 TTGATGACACTTATGATACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 281 CCGAGACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 340
Db 241 CCGAGACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 341 GCTGTGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 400
Db 301 GCTGTGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 401 CGGCGCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 460
Db 361 CGGCGCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 461 GTCCAGAAAGCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 520
Db 421 GTCCAGAAAGCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 521 GCTGTGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 580
Db 481 GCTGTGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 581 CATGGGACATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 640
Db 541 CATGGGACATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 641 GTGCTCTTCCGCAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 700
Db 601 GTGCTCTTCCGCAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 660
QY 701 CGTGTATCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 760
Db 661 CGTGTATCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 761 GAGAACTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 820
Db 721 GAGAACTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 821 GCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 880
Db 781 GCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 840
QY 881 GCAAGCAGATGTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 940
Db 841 GCAAGCAGATGTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 941 TTCCAGAGCTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1000
Db 901 TTCCAGAGCTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 1001 CCGTGTGAGGCTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1060


```

Db      301  GTCCAGGAGAGCCCTGAGCGTGTGAGAGCCCTAGACGCGGCTGCCCGGCAAAACATGTG 360
Qy      521  GCTGGGCCCACTGAGGTCCTGAGCTGTAACCAAGTCTCTTACGCTTTTACCTGGACC 580
Db      361  GCTGGGCCCACTGAGGTCCTGAGCTGTAACCAAGTCTCTTACGCTTTTACCTGGACC 420
Qy      581  CATGGGAGCATGCAACAGCTGTGTGCAACCGCCACCGTATCCTGAGCCTGAGGCTGCC 640
Db      421  CATGGGAGCATGCAACAGCTGTGTGCAACCGCCACCGTATCCTGAGCCTGAGGCTGCC 480
Qy      641  GTGTCTTCCGCCAGATGGCCACCGCCCTGGCCGCACTGTCAACAGCAAGTCTGTGCTG 700
Db      481  GTGTCTTCCGCCAGATGGCCACCGCCCTGGCCGCACTGTCAACAGCAAGTCTGTGCTG 540
Qy      701  CGTATCTCAAGCTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 760
Db      541  CGTATCTCAAGCTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Qy      761  GAGAACCTGGAGGACTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 820
Db      601  GAGAACCTGGAGGACTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Qy      821  GCGTCCCACTGACCTGACCTGAGGACTGAGATTAAGTCAAGCTCAAGGCTTCACTCGG 880
Db      661  GCGTCCCACTGACCTGACCTGAGGACTGAGATTAAGTCAAGCTCAAGGCTTCACTCGG 720
Qy      881  GCAGCCGATGTCTGAGGCTGTGGGCTGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 940
Db      721  GCAGCCGATGTCTGAGGCTGTGGGCTGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Qy      941  TTCAGGACTCGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1000
Db      781  TTCAGGACTCGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Qy      1001  CCTGGAGGCTCTGCGGCTCTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1060
Db      841  CCTGGAGGCTCTGCGGCTCTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Qy      1061  GGTGAAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAAC 1120
Db      901  GGTGAAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAAC 960
Qy      1121  CCTTTCAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCA 1180
Db      961  CCTTTCAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCA 1020
Qy      1181  GGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1234
Db      1021  GGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1074

RESULT 9
US-10-649-156-9
; Sequence 9, Application US/10649156
; Publication No. US20040038346A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. US20040038346A1el Human Protein Kinases and Uses
; FILE REFERENCE: 35800/20996
; CURRENT APPLICATION NUMBER: US/10/649,156
; PRIOR FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/09/799,875
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659,287
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 9
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-649-156-9

Query Match      50.5%; Score 1069.2; DB 13; Length 1074;
Best Local Similarity 99.7%; Pred. No. 6,6e-308;
Matches 1071; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      161  ATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 220
Db      1  ATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
Qy      221  TTGAGATGACAACTTAATACAGAGCGTCCGCTCAGAAACAGAGCTGAGAGTGGAGCCAG 280
Db      61  TTGAGATGACAACTTAATACAGAGCGTCCGCTCAGAAACAGAGCTGAGAGTGGAGCCAG 120
Qy      281  CCCAGAGTGGGCGGCGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 340
Db      121  CCCAGAGTGGGCGGCGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 180
Qy      341  GCTGTGGGCACTGCTCCGCTCTTGGGCGCTATGTCTCTCTGAGGCGGAGAGGAGGAGG 400
Db      181  GCTGTGGGCACTGCTCCGCTCTTGGGCGCTATGTCTCTCTGAGGCGGAGAGGAGGAGG 240
Qy      401  CGGAGCTTACGAGGCGCTGCACTGCGCTTACAGGCACTGATTAAGTCAAGGAGTGTAC 460
Db      241  CGGAGCTTACGAGGCGCTGCACTGCGCTTACAGGCACTGATTAAGTCAAGGAGTGTAC 300
Qy      461  GTTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 520
Db      301  GTTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Qy      521  GCTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 580
Db      361  GCTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
Qy      581  CATGGGAGCATGCAACAGCTGTGTGCAACGCGCACCGTATCCCTGAGCCTGAGGCTGCC 640
Db      421  CATGGGAGCATGCAACAGCTGTGTGCAACGCGCACCGTATCCCTGAGCCTGAGGCTGCC 480
Qy      641  GTGTCTTCCGCCAGATGGCCACCGCCCTGGCCGCACTGTCAACAGCAAGTCTGTGCTG 700
Db      481  GTGTCTTCCGCCAGATGGCCACCGCCCTGGCCGCACTGTCAACAGCAAGTCTGTGCTG 540
Qy      701  CGTATCTCAAGCTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 760
Db      541  CGTATCTCAAGCTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Qy      761  GAGAACCTGGAGGACTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 820
Db      601  GAGAACCTGGAGGACTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Qy      821  GCGTCCCACTGACCTGACCTGAGGACTGAGATTAAGTCAAGCTCAAGGCTTCACTCGG 880
Db      661  GCGTCCCACTGACCTGACCTGAGGACTGAGATTAAGTCAAGCTCAAGGCTTCACTCGG 720
Qy      881  GCAGCCGATGTCTGAGGCTGTGGGCTGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 940
Db      721  GCAGCCGATGTCTGAGGCTGTGGGCTGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Qy      941  TTCAGGACTCGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1000
Db      781  TTCAGGACTCGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Qy      1001  CCTGGAGGCTCTGCGGCTCTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1060
Db      841  CCTGGAGGCTCTGCGGCTCTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Qy      1061  GGTGAAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAAC 1120
Db      901  GGTGAAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAACAGGCTCAAC 960

```


QY 1121 CCTTAGCCCAACCCGATCCATCTCTGGAGGCTGCCAGAGTGTCCCTGATGACTG 1180
 DB 961 CCTTAGCTCCACCCGATCCATCTCTGGAGGCTGCCAGAGTGTCCCTGATGACTG 1020
 QY 1181 GGGCTGACGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1234
 DB 1021 GGGCTGACGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074

RESULT 10

US-09-919-580-580
 ; Sequence 580, Application US/09919580
 ; Patent No. US20020110832A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pyle, Ruth
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Secrist, Heather
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
 ; FILE REFERENCE: 210121.552
 ; CURRENT APPLICATION NUMBER: US/09/919,580
 ; CURRENT FILING DATE: 2001-07-30
 ; NUMBER OF SEQ ID NOS: 934
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 580
 ; LENGTH: 541
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-919-580-580

Query Match 25.2%; Score 533; DB 9; Length 541;
 Best Local Similarity 99.1%; Pred. No. 3,9e-148;

Matches 536; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1181 GGGCTGACGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1240
 DB 1 GGGCGGCTGACGAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
 QY 1241 CACCCCTACACAGCTCAGCTGCAAGTGAATGATTTGGGGGTAGCTCCAAAGCTT 1300
 DB 61 CACCCCTACACAGCTCAGCTGCAAGTGAATGATTTGGGGGTAGCTCCAAAGCTT 120
 QY 1301 CTCCTGCTGTAAGTGAAGCAAGCAAGCTTCAAGTCTCCAGAGAGAGAGAGAGAG 1360
 DB 121 CTCCTGCTGTAAGTGAAGCAAGCAAGCTTCAAGTCTCCAGAGAGAGAGAGAGAG 180
 QY 1361 CTGTGTGAGAGTGTCTGTGTGACATCTGCTTTGTTCCACACATGCAAGTTCCTGCTTG 1420
 DB 181 CTGTGTGAGAGTGTCTGTGTGACATCTGCTTTGTTCCACACATGCAAGTTCCTGCTTG 240
 QY 1421 GGGCTGTAAGTGAAGCAAGCAAGCTTCTGGGTGGGAGTACAGAGAGAGAGAGAG 1480
 DB 241 GGGCTGTAAGTGAAGCAAGCAAGCTTCTGGGTGGGAGTACAGAGAGAGAGAGAGAG 300
 QY 1481 GACATATTCCTGCTGACAGAGATGACAACTGCGATCCTTGAAGCTGACAACTTTTC 1540
 DB 301 GACATATTCCTGCTGACAGAGATGACAACTGCGATCCTTGAAGCTGACAACTTTTC 360
 QY 1541 CATGACCAATGATGATCTGTCTACACTGTGGTACACTTTGTCACATGTGCGGCTTCACATGA 1600
 DB 361 CATGACCAATGATGATCTGTCTACACTGTGGTACACTTTGTCACATGTGCGGCTTCACATGA 420
 QY 1601 TGGCTGAGTCAAGGCAAGCTGCTGTCAGAGCAATCCCTTTCACAAACCAAGCTGCTT 1660
 DB 421 TGGCTGAGTCAAGGCAAGCTGCTGTCAGAGCAATCCCTTTCACAAACCAAGCTGCTT 480
 QY 1661 TTGATCTGTACCTTTTTCAGAGAAAGAGATTCCTGTCGCAAGGCTCCAGGCTTC 1720
 DB 481 TTGATCTGTACCTTTTTCAGAGAAAGAGATTCCTGTCGCAAGGCTCCAGGCTTC 540
 QY 1721 T 1721

DB 541 T 541

RESULT 11

US-09-867-701-10282/c
 ; Sequence 10282, Application US/09867701
 ; Patent No. US20020132237A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aglate, Paul A.
 ; APPLICANT: Jones, Robert
 ; APPLICANT: Harlocker, Susan L.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
 ; FILE REFERENCE: 210121.497
 ; CURRENT APPLICATION NUMBER: US/09/867,701
 ; CURRENT FILING DATE: 2001-05-29
 ; NUMBER OF SEQ ID NOS: 10912
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10282
 ; LENGTH: 426
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-867-701-10282

Query Match 20.1%; Score 426; DB 9; Length 426;
 Best Local Similarity 100.0%; Pred. No. 2.9e-116;

Matches 426; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1679 CAGAGAAAGGAGATATCCCTGTGCCAAAGGCTCCAGGCTTCCTCCAGCACTCAGAGC 1738
 DB 426 CAGAGAAAGGAGATATCCCTGTGCCAAAGGCTCCAGGCTTCCTCCAGCACTCAGAGC 367
 QY 1739 CCAAGCCAGCTCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1798
 DB 366 CCAAGCCAGCTCAGCTGAG 307
 QY 1799 TTCTCTCTCAG 1858
 DB 306 TTCTCTCTCAG 247
 QY 1859 TAGTCTTCTCTAAGTCAAG 1918
 DB 246 TAGTCTTCTCTAAGTCAAG 187
 QY 1919 CTTCTGAGCTGAGCAAG 1978
 DB 186 CTTCTGAGCTGAGCAAG 127
 QY 1979 CTGGAAGTCCAGAGTGGAGCTTCTGAGGAGACAGGATTAGCAGAGGCTGCTGTCGAC 2038
 DB 126 CTGGAAGTCCAGAGTGGAGCTTCTGAGGAGACAGGATTAGCAGAGGCTGCTGTCGAC 67
 QY 2039 ACTCTAGTTTGGATACATGATGATGATTTTACTGTGCTATTAAGAGAGATTA 2098
 DB 66 ACTCTAGTTTGGATACATGATGATGATTTTACTGTGCTATTAAGAGAGATTA 7
 QY 2099 TGAAT 2104
 DB 6 TGAAT 1
 RESULT 12
 US-09-925-301-525
 ; Sequence 525, Application US/09925301
 ; Patent No. US2002052308A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA106
 ; CURRENT APPLICATION NUMBER: US/09/925,301
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05882
 ; PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 525
LENGTH: 562
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (515)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (526)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (557)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-525

Query Match 17.5%; Score 369.4; DB 9; Length 562;
Best Local Similarity 97.2%; Pred. No. 2,4e-99;
Matches 378; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 1 GGAAGCGGCTCCGCGCGGCTCCGCTGCTAGAGACCAGGCGCTGAGCTGGAGCTGGGA 60
DB 161 GGAAGCGGCTCCGCGCGGCTCCGCTGCTAGAGACCAGGCGCTGAGCTGGAGCTGGGA 220
QY 61 TCCGAGCTCGGAGCGAGCGCA-CGGGCGGCGCCCACTGCTGGTGGCTTGAAGCTTGA 119
DB 221 TCCGAGCTCGGAGCGAGCGCA-CGGGCGGCGCCCACTGCTGGTGGCTTGAAGCTTGA 280
QY 120 GCCCGCGCGCGCGCGCGCGCGCGCGCGAGAGCGAGGCGAGATGCGAGCGAGCGCGCTGG 179
DB 281 GCCCGCGCGCGCGCGCGCGCGCGCGCGAGAGCGAGGCGAGATGCGAGCGAGCGCGCTGG 340
QY 180 CTGCTCTGCGGGGTTCTCTCTCCAGAGAGAGCGGTGAGTGGATGACAACTTAGATA 239
DB 341 CTGCTCTGCGGGGTTCTCTCTCCAGAGAGAGCGGTGAGTGGATGACAACTTAGATA 400
QY 240 CCGAGCGCGCGCGCGCGAGAAAGAGAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGG 239
DB 401 CCGAGCGCGCGCGCGCGAGAAAGAGAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGG 460
QY 300 TGTGGCG 359
DB 461 TGTGGCG 520
QY 360 GTCCTGGCG 388
DB 521 GTCCTGGCG 549

RESULT 13

US-09-803-719-519
Sequence 519, Application US/09803719
Publication No. US20030044783A1
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Escobedo, Jaime
APPLICANT: Innis, Michael A.
APPLICANT: Garcia, Pablo Dominguez
APPLICANT: Sudduth-Klinger, Julie
APPLICANT: Reinhard, Christoph
APPLICANT: Giese, Klaus
APPLICANT: Randazzo, Filippo
APPLICANT: Kennedy, Giulia C.
APPLICANT: Pot, David
APPLICANT: Kassam, Altaf
APPLICANT: Lamson, George
APPLICANT: Drmanac, Radoje
APPLICANT: Cirkenjakov, Radomir
APPLICANT: Dickson, Mark

APPLICANT: Drmanac, Snezana
APPLICANT: Labat, Ivan
APPLICANT: Leshkowitz, Dena
APPLICANT: Kita, David
APPLICANT: Garcia, Veronica
APPLICANT: Jones, Lee William
APPLICANT: Stache-Crain, Birgit
TITLE OF INVENTION: Human Genes and Gene Products
FILE REFERENCE: 1624.002
CURRENT APPLICATION NUMBER: US/09/803,719
CURRENT FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/188,609
PRIOR FILING DATE: 2000-03-09
NUMBER OF SEQ ID NOS: 2396
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 519
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapiens
US-09-803-719-519

Query Match 16.7%; Score 354.2; DB 10; Length 396;
Best Local Similarity 96.4%; Pred. No. 7.2e-95;
Matches 373; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 70 CCGCAGAGGCGCA-CGGGCGGCGCCCACTGCTGGTGGCTTGAAGCTTGAAGCTTGAAGCTTGA 128
DB 9 CCGCAGAGGCGCACTTGGCTTTTCCACCTGCTGGTGGCTTGAAGCTTGAAGCTTGAAGCTTGA 68
QY 129 GCGCGCGCGCGCGCGCGCGCGCGCGCGAGAGCGAGTGGATGACAACTTAGATACGAGCGTT 188
DB 69 GCGCGCGCGCGCGCGCGCGCGCGCGCGAGAGCGAGTGGATGACAACTTAGATACGAGCGTT 128
QY 189 CGGCTTCTCTGTCAGAGAAAGCGGTTGAGTGGATGACAACTTAGATACGAGCGTT 248
DB 129 CGGCTTCTCTGTCAGAGAAAGCGGTTGAGTGGATGACAACTTAGATACGAGCGTT 188
QY 249 CCGTCCAGAGAGAGTGGAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGG 308
DB 189 CCGTCCAGAGAGAGTGGAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGG 248
QY 309 TGAAGCCACTGCTGTCAGATCGTGAAGTGGTGGCGCACTGGCTCCCGCTTGGGG 368
DB 249 TGAAGCCACTGCTGTCAGATCGTGAAGTGGTGGCGCACTGGCTCCCGCTTGGGG 308
QY 369 CCTATGCTCTCTGAGCGCGCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 428
DB 309 CCTATGCTCTCTGAGCGCGCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 368
QY 429 CAGGCACTGATGATACCTGCAAGGTGT 455
DB 369 CAGGCACTGATGATACCTGCAAGGTGT 395

RESULT 14

US-09-873-367C-341/C
Sequence 341, Application US/09873367C
Publication No. US20030165839A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
APPLICANT: Soppet, Daniel
APPLICANT: Andress, Gregory
APPLICANT: Augustus, Meena
APPLICANT: Ebner, Reinhard
APPLICANT: Carter, Kenneth
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
FILE REFERENCE: 689290-64
CURRENT APPLICATION NUMBER: US/09/873,367C
CURRENT FILING DATE: 2003-04-29
PRIOR APPLICATION NUMBER: U.S. 60/236,891
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: U.S. 60/236,842

PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: U.S. 60/244,867
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: U.S. 60/245,084
PRIOR FILING DATE: 2000-11-01
NUMBER OF SEQ ID NOS: 1067
SOFTWARE: PatentIn version 3.0
SEQ ID NO 341
LENGTH: 353
TYPE: DNA
ORGANISM: Homo sapiens
US-09-873-367C-341

Query Match 14.9%; Score 315.4; DB 10; Length 353;
Best Local Similarity 95.2%; Pred. No. 2.5e-83;
Matches 336; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 1756 TGGAACTGTGTT-CCACGATCTCTGCTCTTGAATTAAGAGATTCTCTCCAGGCTT 1814
DB 353 TGGAACTGTGTTCCCGACGATCTCTGCTCTTGAATTAAGAGATTCTCTCCAGGCTT 294
QY 1815 AAGCTGGATTTGGGCGCAGATTAAGATTCGAATAGAGCTAGTCTTGTCTAACT 1874
DB 293 AAGCTGGATTTGGGCGCAGATTAAGATTCGAATAGAGCTAGTCTTGTCTAACT 234
QY 1875 CAAGACTGTTCTGAAATGAGGGTCCAGGCTGCAACATGGGGCTTCTGACCTGAGCAG 1934
DB 233 CAAGACTGTTCTGAAATGAGGGTCCAGGCTGCAACATGGGGCTTCTGACCTGAGCAG 174
QY 1935 CAAGTTGAGGAGCAGAGATTAGGCGAGGTCTGCTGTGCGCCACTGCGAAAGTCCAGGT 1994
DB 173 CAAGTTGAGGAGCAGAGATTAGGCGAGGTCTGCTGTGCGCCACTGCGAAAGTCCAGGT 114
QY 1995 GGGACCTCTCTGGGAGCACTTGGGGTCCCAATCCAGGCTCACTCTAGTTTGGAT 2054
DB 113 GGGACCTCTCTGGGAGCACTTGGGGTCCCAATCCAGGCTCACTCTAGTTTGGAT 54
QY 2055 ACCATGATGATGATTTTACCTGCTTAATTAAGAGATTAATAA 2107
DB 53 ACCATGATGATGATTTTACCTGCTTAATTAAGAGATTAATAA 1

RESULT 15
US-10-221-278-116
Sequence 116, Application US/10221278
Publication No. US20040034208A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-045
CURRENT APPLICATION NUMBER: US/10/221,278
CURRENT FILING DATE: 2002-09-06
PRIOR APPLICATION NUMBER: 09/693,267
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/665,363
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 09/616,847
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 09/596,193
PRIOR FILING DATE: 2000-06-17
PRIOR APPLICATION NUMBER: 09/574,454
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: 09/519,705
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 752
SEQ ID NO 116
LENGTH: 1909
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (135) .. (941)
US-10-221-278-116

Query Match 12.2%; Score 258.4; DB 13; Length 1909;
Best Local Similarity 60.3%; Pred. No. 4.8e-66;
Matches 444; Conservative 0; Mismatches 291; Indels 1; Gaps 1;

QY 370 CTATGTCCTCTGAGAGCCCGAGAGGGGCGGCGCTACACGAGGCTGCACTGCGCTAC 429
DB 103 CTACCTGCTGCTGCGCCCTTACCGAGCGGAGCATGTGTCGGGGGCTGTGATCCACAC 162
QY 430 AGGCACTGATATAC-CTGCAAGGTGTACCCCGTCAGAGAACCTGCGCTGTGAGC 488
DB 163 TGGAGCGCGAGGTGCGGCAAGGTGTTCATTAACACTACAGAGCAAAATCAAGGC 222
QY 489 CTTACGCGCGGCTGCGCGCGCAAGCATGTGCTGCGCGCACTAGAGTCTGCGTGTGA 548
DB 223 CTTACATCAGCTGCGCATGCGACAGCAACATTACGTGATGTGGAAGTATCTTGGGG 282
QY 549 CCCAGCTCCTTACGCGCTTTTCACTGCGAACCCATGGGAGACATGACAGCTGTGCGAA 608
DB 283 AAACCAAGGCTATGTCTTCTTTGAGAGAGACTTTGGGACATGCACTCTATGTGCGAA 342
QY 609 CGCGCAACCGTATCCCTGAGGCTGAGGCTGCGCTCTTCCGCAAGTGCACCGCCG 668
DB 343 GCGGAGAGGCTGCGGAGAGAGAGCGCGCGCTCTTCAAGCAGATTGTCTCGCGG 402
QY 669 TGGCGCATGTATCCAGCAGCGTGTGCTGCTGCTGATCTCAAGCTGTGCTGTGTCT 728
DB 403 TGGCCCATGCGCACAGTCAGGCAATGTGTGCGGAGCTGAGGCTTAAAGAGTTCCTT 462
QY 729 TGGCGACCGGTGAG 788
DB 463 TCTTCAGGAG 522
QY 789 CTGGGCAAGATATTCCTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 848
DB 523 AGGGGAGAGATATGTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 582
QY 849 TACTGACCTGACGAGGCTTATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 908
DB 583 TCTTCAACACCACTGAG 642
QY 909 CGCTCTTACCAATGCTGGCGGCGCACTACCCCTTCCAGAGCTGAGAGCTGTCTCTCT 968
DB 643 TGCTCTACACCTTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702
QY 969 TCGGCAAGATCCGCGCGGCGGCTTACGCTGTGCGGAGAGAGAGAGAGAGAGAGAG 1028
DB 703 TCTCCAAATTTGGGCGTGGAGAGTCTGCAATCTGAGCAGATTTCCCGCAAGCAGGT 762
QY 1029 GTCTGTGCTGCTGCTCTTCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1088
DB 763 GCTTCAATGCGACCTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 822
QY 1089 TCCTGACCCCTGAGCT 1104
DB 823 TACTGACCCCTGATT 838

Search completed: August 29, 2004, 14:34:59
Job time : 972 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 29, 2004, 06:33:57 ; Search time 5534 Seconds
(without alignments)
11418.210 Million cell updates/sec

Title: US-10-070-337-16
Perfect score: 2116
Sequence: 1 ggagcgcgccgcgcgcgt.....tatgaataaaaaaaaaa 2116

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_eshba:*
2: em_eshum:*
3: em_estin:*
4: em_estnu:*
5: em_escov:*
6: em_escpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_hiv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1072.2	50.7	1077	29	AY419139 Homo sapi
2	1007.8	47.6	1074	29	AY419140 Pan trogl
3	940.6	44.5	1201	9	AL578892 AL578892
4	902	42.6	942	13	BX393466 BX393466

5	895.2	42.3	1075	13	BX367264 BX367264
6	893	42.2	1051	9	AL525890 AL525890
7	892	42.2	947	13	BX325344 BX325344
8	878.6	41.5	948	13	BX443303 BX443303
9	872.4	41.2	1201	9	AL556690 AL556690
10	868.8	41.1	975	12	BG575275 BG575275
11	862.4	40.8	1001	9	AL562576 AL562576
12	849.8	40.2	873	13	BQ216198 BQ216198
13	846.2	40.0	881	13	BQ941789 BQ941789
14	843.6	39.9	918	13	BQ223497 BQ223497
15	843.4	39.9	945	13	BX353445 BX353445
16	834	39.4	1201	13	BX426485 BX426485
17	832.8	39.4	1201	13	BX445733 BX445733
18	825.8	39.0	929	9	AL522632 AL522632
19	825.4	39.0	944	13	BX363066 BX363066
20	822.6	38.9	1126	13	BX405957 BX405957
21	816.2	38.6	902	13	BX421477 BX421477
22	813.8	38.5	863	12	BG748142 BG748142
23	811.2	38.3	1062	12	BM557568 BM557568
24	807.2	38.1	837	12	B1861766 B1861766
25	802.2	37.9	1076	12	BX367265 BX367265
26	801	37.9	1053	13	BQ224844 BQ224844
27	797.8	37.7	925	13	BX435881 BX435881
28	797.6	37.7	1036	13	BQ217691 BQ217691
29	796.6	37.6	946	13	BX390085 BX390085
30	786.8	37.2	910	12	BG397691 BG397691
31	785.4	37.1	929	13	BUS00619 BUS00619
32	782.2	37.0	851	12	B1909168 B1909168
33	777	36.7	880	12	BG470123 BG470123
34	774.6	36.6	957	13	BQ949069 BQ949069
35	774.4	36.5	891	9	AL555267 AL555267
36	767.4	36.3	989	10	BE792929 BE792929
37	766.6	36.2	810	12	BM045379 BM045379
38	763.8	36.1	932	13	BQ683898 BQ683898
39	751.2	35.5	914	13	BUS00176 BUS00176
40	751	35.5	931	12	BG470193 BG470193
41	742.4	35.1	820	12	B1256421 B1256421
42	738.6	34.9	1002	12	BM804790 BM804790
43	737	34.8	910	9	AL562114 AL562114
44	735.4	34.8	910	12	BG828172 BG828172
45	735.2	34.7	861	9	AL555266 AL555266

ALIGNMENTS

RESULT 1
LOCUS AY419139 1077 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens C20orf97 gene, VIRUAL TRANSCRIPT, partial sequence,
ACCESSION AY419139
VERSION AY419139.1 GI:39775099
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1077)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Clark,A.G., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLES Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1077)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLES Direct Submission

JOURNAL

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
These sequences were made by sequencing genomic exons and ordering
them based on alignment.

COMMENT

FEATURES

Location/Qualifiers
1..1077

/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
<1..>1077
/gene="C20orf97"
/locus_tag="HCM6791"

ORIGIN

Query Match 50.7%; Score 1072.2; DB 29; Length 1077;
Best Local Similarity 99.7%; Pred. No. 6.7e-194;
Matches 1074; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 161 ATGCGAGCCACCCCTCTGGCTGCTCTGCGGGTTCCCTGTCAGAGAAAGCGTTGAG 220
Db 1 ATGCGAGCCACCCCTCTGGCTGCTCTGCGGGTTCCCTGTCAGAGAAAGCGTTGAG 60
QY 221 TTGGATGACAATTAGTATACCGAGCCCTCCGTCCAGAAAGAGCTCGAAGTGGCCCAAG 280
Db 61 TTGGATGACAATTAGTATACCGAGCCCTCCGTCCAGAAAGAGCTCGAAGTGGCCCAAG 120
QY 281 CCCAGACTGCCCCCTGCTGTTGACCCCTGAGCCCACTGCTGTCAGATTCGTCAGACT 340
Db 121 CCCAGACTGCCCCCTGCTGTTGACCCCTGAGCCCACTGCTGTCAGATTCGTCAGACT 180
QY 341 GCTGTGGCCACTGCTCCGCTCTTGGGCTTATGTCCTCTCTGAGCCCGAGAGGCGG 400
Db 181 GCTGTGGCCACTGCTCCGCTCTTGGGCTTATGTCCTCTCTGAGCCCGAGAGGCGG 240
QY 401 CGGAGCCACAGGCGCCCTGCACTCCCTCAACGCACTGATATCTGCAAGTGTACCC 460
Db 241 CGGAGCCACAGGCGCCCTGCACTCCCTCAACGCACTGATATCTGCAAGTGTACCC 300
QY 461 GTCCAGAGAGCCCTGCGCTGCTGAGAGCCCTAGCGCGGCTGCCCGCGAGAGCATATG 520
Db 301 GTCCAGAGAGCCCTGCGCTGCTGAGAGCCCTAGCGCGGCTGCCCGCGAGAGCATATG 360
QY 521 GCTCGGCCCACTGAGTCTGCTGCTGATCCAGACTCTCTCAAGCCTTTTCACTCGGACC 580
Db 361 GCTCGGCCCACTGAGTCTGCTGCTGATCCAGACTCTCTCAAGCCTTTTCACTCGGACC 420
QY 581 CATGGGACATGACAGCTGAGGAGAGCGCGCAACCGATACCTCTAGAGCTGAGCTGCC 640
Db 421 CATGGGACATGACAGCTGAGGAGAGCGCGCAACCGATACCTCTAGAGCTGAGCTGCC 480
QY 641 GTGCTCTTCCGCAATGAGCCACCGGCTGAGGCACTGTCACAGAGAGGTGCTCCCTG 700
Db 481 GTGCTCTTCCGCAATGAGCCACCGGCTGAGGCACTGTCACAGAGAGGTGCTCCCTG 540
QY 701 CGTGAATCTAAGCTGTGTGCTTTGTCTTCGCTGACCGGTGAGAGAGAAAGCTGTGCTG 760
Db 541 CGTGAATCTAAGCTGTGTGCTTTGTCTTCGCTGACCGGTGAGAGAGAAAGCTGTGCTG 600
QY 761 GAGAACCTGAGAGACTCCGCGGTGCTGATGAGGCGCAAGATTTCTCTGTGGAGCAAGAC 820
Db 601 GAGAACCTGAGAGACTCCGCGGTGCTGATGAGGCGCAAGATTTCTCTGTGGAGCAAGAC 660
QY 821 GGTGCGCCAGACTTACGTGGGACCTGAGATCTAGCTCAGGGCTCTATCTCGGGCAAG 880
Db 661 GGTGCGCCAGACTTACGTGGGACCTGAGATCTAGCTCAGGGCTCTATCTCGGGCAAG 720
QY 881 GCAAGCCGATGTGTGAGGCTGGGGCTGGGCTTTTACCATGCTGGCGGCACTAACCC 940
Db 721 GCAAGCCGATGTGTGAGGCTGGGGCTGGGCTTTTACCATGCTGGCGGCACTAACCC 780
QY 941 TTCCAGAGACTGAGAGCTGCTCTGCTCTTTCGCAAGATCCGCGGGGCTTACGCTTG 1000
Db 781 TTCCAGAGACTGAGAGCTGCTCTGCTCTTTCGCAAGATCCGCGGGGCTTACGCTTG 840

```

```

QY 1001 CCTGAGGCTCTTGAGCCCTCCCGCTGCTGTGCTGCTGCTCTCTGTCGAGGCA 1060
Db 841 CCTGAGGCTCTTGAGCCCTCCCGCTGCTGTGCTGCTGCTCTCTCTGTCGAGGCA 900
QY 1061 GCTGAAGGCTCTCAAGGCTCAAGGCTCTCTCTGACCCCTGCTGCGAGAGCCGATG 1120
Db 901 GCTGAAGGCTCTCAAGGCTCAAGGCTCTCTCTGACCCCTGCTGCGAGAGCCGATG 960
QY 1121 CCCTTACCCCAACCCGATCCATCTCTGAGAGGCTCCAGAGGCTGCTGATGACTG 1180
Db 961 CCCTTACCCCAACCCGATCCATCTCTGAGAGGCTCCAGAGGCTGCTGATGACTG 1020
QY 1181 GGGCTGACGAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGTGTCTGTATGCTAG 1237
Db 1021 GGGCTGACGAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGTGTCTGTATGCTAG 1077

```

```

RESULT 2
LOCUS AY419140 1074 bp DNA linear GSS 17-DEC-2003
DEFINITION Pan troglodytes C20orf97 gene, VIRUTAL TRANSCRIPT, partial
ACCESSION AY419140
VERSION AY419140.1 GI:39775100
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
MAMMALIA: Eutheria; Primates; Catarrhini; Hominoidea; Pan.
REFERENCE 1 (bases 1 to 1074)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civeallo,D.R., Lu,F., Murphy,B.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1074)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civeallo,D.R., Lu,F., Murphy,B.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..1074
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>1074
/gene="C20orf97"
/locus_tag="HCM6791"
ORIGIN
Query Match 47.6%; Score 1007.8; DB 29; Length 1074;
Best Local Similarity 95.3%; Pred. No. 1.3e-181;
Matches 1024; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
161 ATGCGAGCCACCCCTCTGGCTGCTCTGCGGGTTCCCTGTCAGAGAAAGCGTTGAG 220
Db 1 ATGCGAGCCACCCCTCTGGCTGCTCTGCGGGTTCCCTGTCAGAGAAAGCGTTGAG 60
QY 221 TTGGATGACAATTAGTATACCGAGCCCTCCGTCCAGAAAGAGCTCGAAGTGGCCCAAG 280
Db 61 TTGGATGACAATTAGTATACCGAGCCCTCCGTCCAGAAAGAGCTCGAAGTGGCCCAAG 120
QY 281 CCCAGACTGCCCCCTGCTGTTGACCCCTGAGCCCACTGCTGTCAGATTCGTCAGACT 340

```

Db	121	CCGAGACTGCCCCCTGCTGCTGTTGCCCCCTGAGGCCACCTAATGCTCTCAATGCTGAACT	180
QY	341	GCTGTGGACATGTGCTTCCTCCGCTCTTGGGACCTATGCTCTCTGTGAGCCCGAGAGGAGG	400
Db	181	GCTGTGGGCAAGGCTCTCCCGTCTTGAGGACCTATGTCTCTCTGTGAGCCCGAGAGGAGG	240
QY	401	CGGAGCTTACAGAGGCTCTGTCACTGCTCTACAGGACCTAGATATACCTGCAAGGTGTACCC	460
Db	241	CGGAGCTTACCGGGGCTCTGTCACTGCTCCCAAGGACGAGGACGAGGATATACCTGTCAAGGTGTACCC	300
QY	461	GTCCAGAGAGGCCCTGGGCGGTGGTGGAGCCCTTACGCGCGGTGGCCCCCGGACAGACATGTG	520
Db	301	GTGCAGAGAGCCCTGGGCGGTGGTGGAGCCCTTACGCGCGGTGGCCCCCGGACAGACATGTG	360
QY	521	GCTCGGAGCCCATGTAGGTCTGTGGCTGTGATACCCAGTCTCTTACGCGCTTTTCACTGTGAGCC	580
Db	361	GCTCGGAGCCCATGTAGGTCTGTGGCTGTGATACCCAGTCTCTTCACTGTGAGCGCTTTC	420
QY	581	CATGGGAGCATGTACAGGCTGTGGTGGAAAGCGCCACCGTATCCTTGAGGCTGTGGCC	640
Db	421	CGCGGGGACATGTACAGGCTGTGGTGGAAAGCGCCGCGCGCATCCTGTGAGCGCTTGGCGCT	480
QY	641	GTGCTCTTCCGCGCAGATGTGGCCACGCGCTGGCGGCACTGTACACAGGACGAGTGTGCTG	700
Db	481	GCGCTTCTCCGCGCAGATGGCCAGCGGCTGGCGGCACTGTACACAGGACGAGTGTGCTG	540
QY	701	CGTATCTCAAGCTGTGTGCTTTGCTTTCTTTCGCTGACCGGTGAGAGAAAGTGTGTGCTG	760
Db	541	CGTATCTCAAGCTGTGTGCTTTGCTTTGCTTCACTGACGCTGTGAGAGGAAAGTGTGTGCTG	600
QY	761	GAGAACTGTGAGGACTCTGCGGTGTGACTGTGGGCCAGATGATTTCCCTGTGGGACAAAGCAC	820
Db	601	GAGAACTGTGAGGACTCTGCGGTGTGACTGTGGGCCAGATGATTTCCCTGTGGGACAAAGCAC	660
QY	821	GCGTGGCCAGCCTACTGTGGGACCTTGATATCACTCAAGGACCTCAATATCTGGGGGCAAG	880
Db	661	GCGTGGCCAGCCTACTGTGGGACCTTGATATCACTCAAGGACCTCAATATCTGGGGGCAAG	720
QY	881	GCAGCGGATGTGTGAGGCTGTGGGCGGTGGGCGCTTTCAACATGCTGTGGCGGACCTAACCC	940
Db	721	GCAGCGGATGTGTGAGGCTGTGGGCGGTGGGCGCTTTCAACATGCTGTGGCGGCGCACTNNNCC	780
QY	941	TTCCAGGACTGTGAGGCTGTGCTGCTTTCCGAGCAATTCGCGCGGGGCTTAAGCTCTG	1000
Db	781	TTCCAGGACTGTGAGGCTGTGCTGCTTTCCGAGCAATTCGCGCGGGGCTTAAGCTCTG	840
QY	1001	CCTGTGAGGCTCTCGGCGCTTGGCGCGGCTGTGCTGTGCTGCTGCTTCCGTGGGAGGCA	1060
Db	841	CCTGTGAGGCTCTCGGCGCTTGGCGCGGCTGTGCTGTGCTGCTTCCGTGGGAGGCA	900
QY	1061	GCTGAACGAGCTCAAGCCCAAGGACATTCCTGTGCAACCTGTGGCTGTGCAAGGACCGGATG	1120
Db	901	GCTGAACGAGCTCAAGCCCAAGGACATTCCTGTGCAACCTGTGGCTGTGCAAGGACCGGATC	960
QY	1121	CCCTTAGCCCAACCCGATCCCATCTCTGGAGAGCTGCGCAAGTGTATCCCTGTATGACTG	1180
Db	961	CCCTTAGCCCAACCCGATCCCATCTCTGGAGAGCTGCGCAAGTGTATCCCTGTATGACTG	1020
QY	1181	GCGCTGAGAGAAAGCCGAGGAAAGAGAGAGGAGACAGAGAAAGTGTCTGTATGGC	1234
Db	1021	GCGCTGAGAGAAAGCCGAGGAAAGAGAGAGGAGACAGAGAAAGTGTCTGTATGGC	1074

RESULT 3	1201 bp	mRNA	linear	EST 01-JUN-2003
AL578892/c				
LOCUS	AL578892			
DEFINITION	AL578892 Homo sapiens HELA CELLS	COT 25--NORMALIZED		Homo sapiens
ACCESSION	AL578892	CDNA clone CS0DK005YC05 3--PRIME,		mRNA sequence.
VERSION	AL578892.2	GI:31317080		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			

REFERENCE
1 (bases 1 to 1201)
AUTHORS
Li, W.-B., Gruber, C., Jesse, J. and Polyes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
On Feb 16, 2001 this sequence version replaced gi:12943405.

Eukaryota, Metazoa: Chordata; Craniata; Vertebrata: Euteleostomi, Mammalia; Eutheria: Primates; Catarrhini; Hominidae; Homo. 1 (Bases 1 to 1201) Li, W.-B., Gruber, C., Jesse, J., and Polayes, D. Full-length cDNA libraries and normalization Unpublished (2001) On Feb 16, 2001 this sequence version replaced gi:12943405.

Contact: Genoscope - Centre National de Sequencage Genoscope - Centre National de Sequencage BP 131 91006 Evry cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10000.f For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSDK005AB03MPLcluster=10000.f>. Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Paraday Avenue genoscope sequence ID : CSDK005AB03MPL.

FEATURES
source

```

1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS00K005C05"
/cell_line="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
ligated with Not I and cloned into the Not I and EcoR V
sites of the pOWSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	44.5%	Score 940.6;	DB 9;	Length 1201;
Best Local Similarity	96.1%	Pred. No. 8e-169;		
Matches 943;	Conservative 24;	Mismatches 13;	Indels 1;	Gaps 1

QY	1096	CCCTGCTGGACGACGACCCTGATGCCCTTAAAGCCCAACCCGATCCCATCTCTGGGAGGC	1155
Db	992	MCCCTKSGCGGAMWAGACCCCKASCCCTTDSITCCACCCCTCTWTCCTCCCTTC - RSGAGGC	934
QY	1156	TGCCAGAGTGGTCCCTGATNGACATGGGGGCTGAGCAAGCCAGAGGAAGAGAGGAGACAG	1215L
Db	933	TGCCAAGTGGTCTCTTATKTKMTTKTGGGCTGAGCAAGCCAGAGGAAGAGGAGAGACAG	874
QY	1216	AGAGTGGTCTGTATGGCTAGACCAACCTACTACACGCTCAGGCTGCCAACAGTGAAT	1275L
Db	873	AGAAATGTTCTGTATGGCTAGACCAACCTACTACACGCTCAGGCTGCCAACAGTGAAT	814
QY	1276	GAGTTGGGGGTAAGCTCCAAAGCCTTCTCTGCTCTGAATGAGCCAAACCTTCACTGTC	1335L
Db	813	GAGTTGGGGGTAAGCTCCAAAGCCTTCTCTGCTCTGAATGAGCCAAACCTTCACTGTC	754
QY	1336	TTCAGAGGGAGAAAGGCAGAAAGCCTGTGTGGAAATGGCTGTATGACATCTGCTTGT	1395L
Db	753	TTCAGAGGGAGAAAGGCAGAAAGCCTGTGTGGAAATGGCTGTATGACATCTGCTTGT	694
QY	1396	TCCACACATGAGTCTCTGCTTGGGATCTTATCAGGGTGCACAAACCCTCTTCGATGC	1455L
Db	693	TCCACACATGAGTCTCTGCTTGGGATCTTATCAGGGTGCACAAACCCTCTTCGATGC	634
QY	1456	TGGGATTAAGAGTGTAGCCAAAGAGACAAATATCTCTGCTCAACAGATGACAAACTGG	1515L
Db	633	TGGGATTAAGAGTGTAGCCAAAGAGACAAATATCTCTGCTCAACAGATGACAAACTGG	574
QY	1516	GATCTTTAGCTACAAACCTTTTCATATACATAGTACTGTCTTACATGGGATCACT	1573L
Db	573	CATCTTTAGCTACAAACCTTTTCATATACATAGTACTGTCTTACATGGGATCACT	514
QY	1576	TTTBTACAGTGTGGGCTCAACATGATGTGGTCTCAGGCACTCTGTCTCAAGGACAAATC	1633L
Db	513	TTTBTACAGTGTGGGCTCAACATGATGTGGTCTCAGGCACTCTGTCTCAAGGACAAATC	454

QY 1636 CTTTACAAACAAACAGCTGCTTGTATCTTGTACCTTTTCAGAGAAAGGAGAT 1695
 Db 453 CTTTACAAACAAACAGCTGCTTGTATCTTGTACCTTTTCAGAGAAAGGAGAT 1695
 QY 1696 CTTTGTGCAAAAGGCTTCCAGGCTTCTCCCTGCAACTAGAGACCCAGGCTCAGTC 1755
 Db 393 CTTTGTGCAAAAGGCTTCCAGGCTTCTCCCTGCAACTAGAGACCCAGGCTCAGTC 1755
 QY 1756 TGGAGACTGTCTCCAGCATCTGTCTGTATAGAGATTTCTTCCAGGCTTA 1815
 Db 333 TGGAGACTGTCTCCAGCATCTGTCTGTATAGAGATTTCTTCCAGGCTTA 1815
 QY 1816 AGCTGGATTTGGGCCAGAGATTAAGATTCAGATAGAGCTAGTCTTGTCTAATC 1875
 Db 273 AGCTGGATTTGGGCCAGAGATTAAGATTCAGATAGAGCTAGTCTTGTCTAATC 1875
 QY 1876 AAGACTGTCTGGAATAGAGGCTTCCAGGCTTCTCAACATGGGCTTGTAGAGAC 1935
 Db 213 AAGACTGTCTGGAATAGAGGCTTCCAGGCTTCTCAACATGGGCTTGTAGAGAC 1935
 QY 1936 AAGTTGAGGACAGATTAAGAGGCTTCTCCCTGCAACTAGAGGCTTGTAGAGAC 1995
 Db 153 AAGTTGAGGACAGATTAAGAGGCTTCTCCCTGCAACTAGAGGCTTGTAGAGAC 1995
 QY 1996 GGAATCTTCTGGGACACTTGGGCTTCCAAATCCAGGCTCAGTCTAGGTTTGGATA 2055
 Db 93 GGAATCTTCTGGGACACTTGGGCTTCCAAATCCAGGCTCAGTCTAGGTTTGGATA 2055
 QY 2056 CCATGATATGTATGTATACC 2076
 Db 33 CCATGATATGTATGTATACC 13

RESULT 4
 LOCUS BX393466/c
 DEFINITION BX393466 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
 CDNA clone CS0DB003YPI9 3-PRIME, mRNA sequence.
 ACCESSION BX393466
 VERSION BX393466.1 GI:30620023
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 L1, W.B., Gruber, C., Jesse, J. and Polyes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 10000.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DB003CHI1ONP1&cluster=10000.f. Contact :
 Feng Liang Email: fliang@life.technogen.com URL :
 http://fulllength.invitrogen.com/invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DB003CHI1ONP1.
 Location/Qualifiers

FEATURES

source

1. 942
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DB003YPI9"
 /issue_type="NEUROBLASTOMA COT 10-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V

ORIGIN sites of the PCMVSPORT 6 vector. Library was normalized."

Query Match 42.6%; Score 902; DB 13; Length 942;
 Best Local Similarity 97.0%; Pred. No. 1,8e-161;
 Matches 906; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

QY 1147 CTGGAGAGCTGCCAGAGTGTCTCCCTGATGACTGGGGCTGAGAGCAAGCCAGAGAGAGA 1206
 Db 942 CTGGAGAGCTGCCAGAGTGTCTCCCTGATGACTGGGGCTGAGAGCAAGCCAGAGAGAGA 1206
 QY 1207 GGGAGACAGAGAGTGTCTGATAGTGTAGAGCAACCTTATACAGCTCAGCTGCCAA 1266
 Db 882 GGGAGACAGAGAGTGTCTGATAGTGTAGAGCAACCTTATACAGCTCAGCTGCCAA 1266
 QY 1267 CAGTGATAGATTGGGGGTAGCTCCAGAGCTTCTCCTGCTTGAATAGGCAAAAC 1326
 Db 822 CAGTGATAGATTGGGGGTAGCTCCAGAGCTTCTCCTGCTTGAATAGGCAAAAC 1326
 QY 1327 TTGAGTGTCTCCAG 1386
 Db 762 TTGAGTGTCTCCAG 1386
 QY 1387 CTGCTTGTCTCCAG 1446
 Db 702 CTGCTTGTCTCCAG 1446
 QY 1447 TCTGGTGTCTGGAGAGTACAGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1506
 Db 642 TCTGGTGTCTGGAGAGTACAGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1506
 QY 1507 ACAAGCTGAGATCTTGAAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1566
 Db 582 ACAAGCTGAGATCTTGAAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1566
 QY 1567 GGGTACACTTTGACAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1626
 Db 522 GGGTACACTTTGACAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1626
 QY 1627 AGAGCAATCCCTTTCACAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1686
 Db 462 AGAGCAATCCCTTTCACAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1686
 QY 1687 GGGAGATTCCTGAG 1746
 Db 402 GGGAGATTCCTGAG 1746
 QY 1747 AGCTCACTGGGAACTGTGTTCCAGAGATCTGTCTCTGTATAGAGATTTCCTT 1806
 Db 342 AGCTCACTGGGAACTGTGTTCCAGAGATCTGTCTCTGTATAGAGATTTCCTT 1806
 QY 1807 CAGAGCTTAAGCTGGGATTTGGGCCAGAGATTAAGATTCAGAGCTAGTCTT 1866
 Db 282 CAGAGCTTAAGCTGGGATTTGGGCCAGAGATTAAGATTCAGAGCTAGTCTT 1866
 QY 1867 GGTCTACTAAGAGTGTGTGAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1926
 Db 222 GGTCTACTAAGAGTGTGTGAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1926
 QY 1927 CTGAGACCAAGAGTGTGAG 1986
 Db 162 CTGAGACCAAGAGTGTGAG 1986
 QY 1987 TCCAGAGTGGAGCTTCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2046
 Db 102 TCCAGAGTGGAGCTTCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2046
 QY 2047 TTTTGATACCAATGATATGATATGATATGATATGATATGATATGATATGATATGATAT 2080
 Db 42 TTTTGATACCAATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATAT 2080

RESULT 5

source 1. 1051

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS00C013Y08"

/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 42.2%; Score 893; DB 9; Length 1051;

Best Local Similarity 95.7%; Pred. No. 9.3e-160;

Matches 908; Conservative 25; Mismatches 14; Indels 2; Gaps 2;

1122 CCTTACCCCAACCCGATCCCATCTCTGGAGGCTGCCAGGTGCTCTGATGACTGG 1181

949 CCTTACCTCCACCCGATCCCATCTCTGGAGGCTGCCAGGTGCTCTGATGACTGG 890

1132 GCTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1241

889 GAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 830

1242 ACCCTACTACAGCTCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1301

829 ACCCTACTACAGCTCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 770

1302 TCTTGCCTCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1361

769 TCTTGCCTCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 710

1362 TGTGTGAGT 1421

709 TGTGTGAGT 650

1422 GTGCTTATCAGGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1481

649 GTGCTTATCAGGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 590

1482 ACAATATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1541

589 ACAATATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 530

1542 ATGACCATAGGTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1601

529 ATGACCATAGGTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 470

1602 GCTGTGCTGAGGACCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1661

469 GCTGTGCTGAGGACCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 410

1662 TGTATCTTGTACCTTTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1721

409 TGTATCTTGTACCTTTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 350

1722 CCCCCTGAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1781

349 CCMCMSCAATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 290

1782 TCTCTTGTATTAAGATTTCTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1841

289 TCTCTTGTATTAAGATTTCTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 230

1842 AATCCAACTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1901

229 AATCCAACTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 170

1902 GCTTGTCAACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1961

169 GCTTGTCAACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 110

QY 1962 GTCCTGCTGT 2021

DB 109 GCTGTGCTGT 51

QY 2022 CACAAATCCGAGT 2070

DB 50 BAAATTCGAGT 3

RESULT 7

BX325344/C

LOCUS

BX325344 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED

DEFINITION

Homo sapiens cDNA clone CS00L004YF04 3-PRIME, mRNA sequence.

ACCESSION

BX325344.1 GI:30336457

VERSION

EST.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.

AUTHORS

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Contact: Genoscope

FEATURES

Genoscope - Centre National de Sequencage

source

BP 191 91006 EVRY cedex - France

location/Qualifiers

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

1. 947

Library was constructed by Life Technologies, a division of

/organism="Homo sapiens"

Invitrogen Corporation 1600

/mol_type="mRNA"

Paradey Avenue Genoscope sequence ID: CS0AL004D02NPL

/db_xref="taxon:9606"

http://fulllength.invitrogen.com/

/clone="CS00L004YF04"

http://fulllength.invitrogen.com/

/cell_type="B CELLS (RAMOS CELL LINE)"

http://fulllength.invitrogen.com/

/clone_lib="RAMOS CELL LINE"

http://fulllength.invitrogen.com/

/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

http://fulllength.invitrogen.com/

Query Match

42.2%; Score 892; DB 13; Length 947;

Best Local Similarity

96.4%; Pred. No. 1.4e-159;

Matches 907; Conservative 22; Mismatches 10; Indels 2; Gaps 2;

1142 CATCTGTGGAGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1201

DB 947 CTTCTGTGGAGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 888

QY 1202 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1261

DB 887 RAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 828

QY 1262 GCCAAGTGTAGT 1321

DB 827 GCCAAGTGTAGT 768

QY 1322 AAACCTTCACTGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1381

DB 767 AAACCTTCACTGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 708

QY		1382	CACATCTGTTTGTTCACACACATGCAGTCTCCTTGGGGCTATTACAGGCCAACG	1441
Ddb		707	CACATCTGTTTGTTCMACACAATSAFTCTCTTGTTGGGCTTAATAGGTCCAAAC	648
QY		1442	CGTGTTCCTGGGTCGGGAGTACAGCAGTAGGACCAAGAATATTCCTGTCTACAG	1501
Ddb		647	CGTGTTCCTGGGTCGGGAGTACAGCAGTAGGACCAAGAATATTCCTGTCTACAG	588
QY		1502	AGATGACAAACTGGCATCCCTTGAGCTGACCAACATTTTCATGACCATAAGTACTGTCT	1561
Ddb		587	AATATGACAAACTGGCATCCCTTGAGCTGACCAACATTTTCATGACCATAAGTACTGTCT	528
QY		1562	AACCTGGGTAACACTTTGTATCCAATGTCTGCGCTTCACATGATGCTGGGTCACGCACTCT	1621
Ddb		527	ACACTGGGTAACACTTTGTATCCAATGTCTGCGCTTCACATGATGCTGGGTCACGCACTCT	469
QY		1622	GTCOCAGGACAAATCCCTTTCACAAACAAACCAAGCTCCTTTGATCTTGATCCTTTTCAG	1681
Ddb		468	GTCOCAGGACAAATCCCTTTCACAAACAAACCAAGCTCCTTTGATCTTGATCCTTTTCAG	409
QY		1682	AGAAAGGAGGTATCCCTGTGCCAAAGGCTCCAGGCTCTCCCTGTGCACTCAGACCCA	1741
Ddb		408	AGAAAGGAGGTATCCCTGTGCCAAAGGCTCCAGGCTCTCCCTGTGCACTCAGACCCA	349
QY		1742	AGCCCAAGCTCACCTGTGGGAATGTGTTCACAGCATCTGTCTCCTTGATTAAGATTC	1801
Ddb		348	AGCCCAAGCTCACCTGTGGGAATGTGTTCACAGCATCTGTCTCCTTGATTAAGATTC	289
QY		1802	TCTTTCAGGCTCAAGCTGTGGGATTTGGGCGACAGATTAAGATCAACTATGAGCTAG	1861
Ddb		288	TCTTTCAGGCTCAAGCTGTGGGATTTGGGCGACAGATTAAGATCAACTATGAGCTAG	229
QY		1862	TTCTTGTCTAAGCTCAAGCTGTTCGTGAATGAGGGTCCAGGCTGTCAACCATGGGCTT	1921
Ddb		228	TTCTTGTCTAAGCTCAAGCTGTTCGTGAATGAGGGTCCAGGCTGTCAACCATGGGCTT	169
QY		1922	CTGACTGTGACCAAGGTTGAGGGAACAGATTTAGSCAGAGGTCTGTCTGTGGGCACCTG	1981
Ddb		168	CTGACTGTGACCAAGGTTGAGGGAACAGATTTAGSCAGAGGTCTGTCTGTGGGCACCTG	109
QY		1982	GAAGTCCCAGGTGGGACTCTTTCGGGGGACACTTGGGGTCCACATCCCAAGTCCATCT	2041
Ddb		108	GAAGTCCCAGGTGGGACTCTTTCGGGGGACACTTGGGGTCCACATCCCAAGTCCATCT	49
QY		2042	CTAGTTTTGGATAC-CATGAGTATGTATGTATTCCTGTGC	2081
Ddb		48	CTAGTTTTGGATACNONCATGGGTATGTATGTITMBKSGTGS	8
RESULT 8				
BX443303		948 bp	mRNA	linear EST 22-MAY-2003
LOCUS				
DEFINITION			BX443303 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA	
ACCESSION			clone CSODG007YF05 5-PRIME, mRNA sequence.	
VERSION			BX443303	
KEYWORDS			BX443303.1 GI:31018564	
SOURCE			EST.	
ORGANISM			Homo sapiens (human)	
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE			1 (bases 1 to 948)	
JOURNAL			Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
COMMENT			Full-length cDNA libraries and normalization	
			Unpublished (2001)	
			Contact: Genoscope	
			Genoscope - Centre National de Sequencage	
			BP 191 91006 EVRY cedex - France	
			Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr	
			Library was constructed by Life Technologies, a division of	
			Invitrogen. This sequence belongs to sequence cluster 10000.f For	
			more information about this cluster, see	
			http://www.genoscope.cns.fr/	

FEATURES					
SOURCE					
					c91-bin/cluster.cgi?seq=CS0DG007CC03QPL&cluster=10000.f. Contact : Peng Liang Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DG007CC03QPL.
					Location/Qualifiers
					1. .948
					/organism="Homo sapiens"
					/mol_type="mRNA"
					/db_xref="taxon:9606"
					/clone="CS0DG007YF05"
					/tissue_type="B CELLS (RAMOS CELL LINE)"
					/cell_line="RAMOS CELL LINE"
					/clone_id="Homo sapiens B CELLS (RAMOS CELL LINE)"
					/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
ORIGIN					
Query Match	41.5%;	Score 878.6;	DB 13;	Length 948;	
Best Local Similarity	98.0%;	Pred. No. 5,1e-157;			
Matches	917;	Conservative	2;	Mismatches 14;	Indels 3;
				Gaps	3;
Db	21	CGCGTGTAGACCCGCGCAGGGCTGAGAGCTGGAGTCGCCGAGTGCCGACGACG	80		
	15	CCGGTCCGGAATTCGCCGGATGAGCTGAGAGCTGGAGTCCGAGTCCGACGACG	74		
Qy	81	CA-CGGGCGGCGCCACCTGCGTGGTGCCTTGAGAGCTGAGACC CGCGCGCGCGCGGCGC	139		
Db	75	CAGNNGNCCTGGCCACTGCTGGTGGTCCCTGAGAGCTCTGAGCCCCGCGGCGCGCGCC	134		
Qy	140	CACGCGAACAACGCGCGGCGAGATGCGAACCAACCCCTCTGCTCTCTGCGGATTCCCTG	199		
Db	135	CACGCGAACAACGCGGCGGCGAGATGCGAACCAACCCCTCTGCTCTCTGCGGATTCCCTG	194		
Qy	200	TCCAGAGAAGACGGCTGTGAGTTGGATGACAACTTAGTATCCGACACGTCCTGTCAGAAA	259		
Db	195	TCCAAGAGAAGACGGCTGTGAGTTGGATGACAACTTAGTATCCGACACGTCCTGTCAGAAA	254		
Qy	260	CGAGCTCGAAGTGGGCGCCGAGCCGACGACCTGCCCTGCTTTGACCCCTGAGCCACT	319		
Db	255	CGAGCTCGAAGTGGGCGCCGAGCCGACGACCTGCCCTGCTTTGACCCCTGAGCCACT	314		
Qy	320	ACTGCTCGAAGTCGTGCAACTGCTGTGGCACTGCTCCGTCCTTGGGCGCTATGTCTTC	379		
Db	315	ACTGCTCGAAGTCGTGCAACTGCTGTGGCACTGCTCCGTCCTTGGGCGCTATGTCTTC	374		
Qy	380	CTGGAGCCCCGAGAGGGCGCGGCGGCGCTTACAGAGCCCTGACATGCCCTTACAGCACTAG	439		
Db	375	CTGGAGCCCCGAGAGGGCGCGGCGGCGCTTACAGAGCCCTGACATGCCCTTACAGCACTAG	434		
Qy	440	TATACCTGCAAGGTATACCCCGTCCAGAAAGCCCTGGCGTGTGAGAGCCCTTACGCGCG	499		
Db	435	TATACCTGCAAGGTATACCCCGTCCAGAAAGCCCTGGCGTGTGAGAGCCCTTACGCGCG	493		
Qy	500	CTGCCCCCGCACAGCATGTGCTCGGACCACTGAGCTCTTGCTGTATCCAGCTCTTC	559		
Db	494	CTGCCCCCGCACAGCATGTGCTCGGACCACTGAGCTCTTGCTGTATCCAGCTCTTC	553		
Qy	560	TAGCGCTTTTCACTGGAGCCCATTTGGAGCATGCAAGCGCTGTGGAGAGCGCCACCGT	619		
Db	554	TAGCGCTTTTCACTGGAGCCCATTTGGAGCATGCAAGCGCTGTGGAGAGCGCCACCGT	612		
Qy	620	ATCCCTGAGCTGAGGCTGCGTGTCTTCCGCGAGATGGCCACCGCCCTGGCGCACTGT	679		
Db	613	ATCCCTGAGCTGAGGCTGCGTGTCTTCCGCGAGATGGCCACCGCCCTGGCGCACTGT	672		
Qy	680	CACGACGAGCTGTGATCTGCGGTGATCTCAAGCTGTGTGCTTTGTCTTCGCTGACCGT	733		
Db	673	CACGACGAGCTGTGATCTGCGGTGATCTCAAGCTGTGTGCTTTGTCTTCGCTGACCGT	732		
Qy	740	GAGAGAGAAGAGCTGTGCTGGAGAACCTGAGAGACTCTCGCGGTGACTGGGCGAGAT	795		

Db 733 GAGAGGAAGAACCTGGTGTCTGGAGAACTCGAGAGACTCTCGGTGCTGACTGGCCAGAT 792

QY 800 GATTCCCTGTGGGACAAGCAGCGGCGCCAGCCTACGTGGGACTGAGTATCTAGCTCA 859

Db 793 GATTCCCTGTGGGACAAAGCACCGGTGCCAGCCCTACGTGGGACCTGAGTACTACACTCA 852

QY 860 CGGCGCTCATTAATCTCGGGCAGAGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTACCC 919

Db 853 CGGCGCTCATTAATCTCGGGCAGAGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTACCC 912

QY 920 ATGCTGGCGGGCCATTACCCCTTCCAGAGCACTGGAG 955

Db 913 ATGCTGGCGGGCCATTACCCCTTCCAGAGCACTGGAG 948

RESULT	9
AL556690	
LOCUS	
DEFINITION	1201 bp mRNA linear EST 31-MAY-2003
ACCESSION	AL556690 Homo sapiens HELA CELLS COR 25-NORMALIZED Homo sapiens
VERSION	CNNA clone CSDDR005YC05 5'-PRIME, mRNA sequence.
KEYWORDS	AL556690 GI:31278491
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Homo sapiens

```

FEATURES
source      1. .1201
            Location/Qualifiers
            ID : CS00NR005AB03Q0P1.
            sequence

```

ORIGIN

Query Match	41.2%;	Score 872.4;	DB 9;	Length 1201;
Best Local Similarity	95.8%;	Pred. No. 7.9e-15;		
Matches 909; Conservative	15;	Mismatches 22;	Indels 3.	Gaps 3.

39 TAGGGCTGAGAGCTGGGGTCTGGGATTCCTCGAGCTGGGACGACGCGACG-66GCGGACCTACT 97
63 CCGGATTGGAGAGCTGGGGCTGGGATTCCTCGAGCTGGGACGACGCGAAGGNGNCCGGGCCACT 122
98 GCTGGTCCCTGGAGAGGCTCTGAGCCCCCGGCGCGCCCGGGGCCCAAGCGGAATGACGCGGCG 157
123 GCTGGTGGCCCTGGAGAGGCTCTGAGCCCCCGGCGCGCCCGGGGCCCAAGCGGAATGACGCGGCG 182
158 GAGATCGAGACACCCCTCTGGCTGCTCTGGGGGATTCCTGTCCAGAGAGAGCGGTTG 217

Db	183	GAGATGGAGCACCACCCCTCTGGCTGCTCTCGCGGGTTCCCTGTCCAGGAAGAGCGGTTG	242
QY	218	GAGTGGATGACAACATTAGATACCGAGCGTCCCGTCCAGAAAGAGCTCGAGTGGGCC	277
Db	243	GAGTTGAGATGACAACATTAGATACCGAGCGTCCCGTCCAGAAAGAGCTCGAGTGGGCC	302
QY	278	CAGCCGAGATGCCCCCTGCTGCTGGGCCCTGAGCCCACTATCTGTCCAGATTCGTGCA	337
Db	303	CAGCCGAGATGCCCCCTGCTGCTGGGCCCTGAGCCCACTATCTGTCCAGATTCGTGCA	362
QY	338	ACTGCTGTGGCCACTGCTCTCCGCTCTTGGGCCCTATATGCTCTCTGAGAGCCGAGAGGCG	397
Db	363	ACTGCTGTGGCCACTGCTCTCCGCTCTTGGGCCCTATATGCTCTCTGAGAGCCGAGAGGCG	422
QY	398	GGCGGGGGCTTACAGGGCCCTGCACTGCTCCCTACAGGCACTGAGTATACCTGCAAGTGTAC	457
Db	423	GGCGGGGGCTTACAGGGCCCTGCACTGCTCCCTACAGGCACTGAGTATACCTGCAAGTGTAC	482
QY	458	CCCGTTCAGAGAACCTTGCGCGGTGCGAGGCCCTACGCGGCGCTGGCCCCGACAGCAT	517
Db	483	CCCGTTCAGAGAACCTTGCGCGGTGCGAGGCCCTACGCGGCGCTGGCCCCGACAGCAT	542
QY	518	GTGGCTTCGCGCCCACTGAGGTCCTGTGCTGGTATCCCAAGTCTCTTCAACGCTTTTCACTGG	577
Db	543	GTGGCTTCGCGCCCACTGAGGTCCTGTGCTGGTATCCCAAGTCTCTTCAACGCTTTTCACTGG	602
QY	578	ACCATTGGGGACATGCAACAAGCTGTGGTGGAAACGCGCAACGTTATCCCTGAGGCTAGGCT	637
Db	603	ACCATTGGGGACATGCAACAAGCTGTGGTGGAAACGCGCAACGTTATCCCTGAGGCTAGGCT	661
QY	638	GGCGGTCCTTCCGCGCAATGAGCCACCGCCCTTGAGCGCACTGTCAACGACAGGCTGTGTC	697
Db	662	GGCGGTCCTTCCGCGCAATGAGCCACCGCCCTTGAGCGCACTGTCAACGACAGGCTGTGTC	721
QY	698	CTGGGTGATCTCAAGCTGTGTGCGCTTTGTCTTGGCTGACCGGTGAGAGAGAAAGTGGTG	757
Db	722	CTGGGTGATCTCAAGCTGTGTGCGCTTTGTCTTGGCTGACCGGTGAGAGAGAAAGTGGTG	781
QY	758	CTGAGAAACCTGAGAGACTCTCGGTGCTGACTGAGGCAGATGATTCCTGTGGGACAAAG	817
Db	782	CTGAGAAACCTGAGAGACTCTCGGTGCTGACTGAGGCAGATGATTCCTGTGGGACAAAG	841
QY	818	CAGCGGTGCCAGCTTACGTGTGGAACTTGAGATACCTGAGCTCACGCGGCTCTTACTCGGGC	877
Db	842	CAGCGGTGCCAGCTTACGTGTGGAACTTGAGATACCTGAGCTCACGCGGCTCTTACTCGGGC	901
QY	878	AAGCAGCCGATCTGAGAGCTGTGGGCTGTGGCGCTTTCACCAATGTGGCGGCACTAC	937
Db	902	AAGCAGCCGATCTGAGAGCTGTGGGCTGTGGCGCTTTCACCAATGTGGCGGCACTAC	961
QY	938	CCCTTTCAGAGACTCGAGCGTGTCTGCTCTCTTGGGCAAGATCCGCGCG	986
Db	962	CCCTTTCAGAGACTCGTGTGTCTTCTCTCTTTCGMAAATATCGCGGCGG	1009

RESULT 10	975 bp	mRNA	linear	EST 10-APR-2001
LOCUS	BG575275			
DEFINITION	BG575275.1 NIH-MGC_87 Homo sapiens cDNA clone IMAGE:4706575 5',			
ACCESSION	mRNA sequence.			
VERSION	BG575275			
KEYWORDS	BG575275.1 GI:13582928			
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1 (bases 1 to 975)			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D.			

Email: c9abbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULNL at:
 http://image.jnl.gov
 Plate: LHAM0575 row: a column: 08
 High quality sequence stop: 891.
 Location/Qualifiers

FEATURES

source

1. 975
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4706575"
 /issue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 87"
 /note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.383 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC library."

ORIGIN

Query Match 41.1%; Score 868.8; DB 12; Length 975;
 Best Local Similarity 96.2%; Pred. No. 3.8e-155;
 Matches 922; Conservative 0; Mismatches 32; Indels 4; Gaps 3;

825 GCCAGCCTGCTGGGAGCTGAGATCTCAGTCAAGGAGCTCATCTCGGAGCAAGCAG 884
 1 GCCAGCCTGCTGGGAGCTGAGATCTCAGTCAAGGAGCTCATCTCGGAGCAAGCAG 60
 885 CCGATGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 944
 61 CCGATGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 120
 945 AGAAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1004
 121 AGAAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 180
 945 AGAAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1004
 1005 CAGGCTCTCGGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1064
 181 CAGGCTCTCGGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 240
 1065 AACGCTCAGAGCCAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1124
 241 AACGCTCAGAGCCAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 300
 1125 TAGCCCTCAGAGCCAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1184
 301 TAGCCCTCAGAGCCAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 360
 1185 TGAAGCAAGCCAG 1244
 361 TGAAGCAAGCCAG 420
 1245 CTATACAGCGCTCAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1304
 421 CTATACAGCGCTCAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 480
 1305 TGCCCTGAGAGCCAG 1364
 481 TGCCCTGAGAGCCAG 540
 1365 GTGAGAGTGTGTGTATACATCTGCTTGTGTATACATCTGCTTGTGTATACATCTGCTTGTGT 1424
 541 GTGAGAGTGTGTGTATACATCTGCTTGTGTATACATCTGCTTGTGTATACATCTGCTTGTGT 600
 1425 CTATACAGTGTGTGTATACATCTGCTTGTGTATACATCTGCTTGTGTATACATCTGCTTGTGT 1484
 601 CTATACAGTGTGTGTATACATCTGCTTGTGTATACATCTGCTTGTGTATACATCTGCTTGTGT 660

QY 1485 ATATTCCTGCTGACAGAGATGACAAACTGGATCCTTGGAGTGCACACATTTTCATG 1544
 DB 661 ATATTCCTGCTGACAGAGATGACAAACTGGATCCTTGGAGTGCACACATTTTCATG 720
 QY 1545 ACCATAGGTGACCTGCTGACATCTGAGTACATTTT--GTACAGATGTGGCTTCACTGATG 1602
 DB 721 ACCATAGGTGACCTGCTGACATCTGAGTACATTTTGTACCAAGTGTGGCTTCACTGATG 780
 QY 1603 CTGTGTCTGAGGACCTGCTGTCAGAGCAATCCTTTTCACAAACCAACAGCTGCTTT 1662
 DB 781 CTGTGTCTGAGGACCTGCTGTCAGAGCAATCCTTTTCACAAACCAACAGCTGCTTT 840
 QY 1663 -GTATCTGTACCTTTTC-AGAGAAAGAGAGTATCCTGTGTCACAAAGGCTCCAGGCTC 1720
 DB 841 GGTATCTGTACCTTTTCAGAGTAAGGAGAGAAATCCCTGGGCAAGGCTTCCAGGCTC 900
 QY 1721 TCCCTGCAACCTCAGAGACCAAGGACCTGCTGAGTGTGTTTCCAGCATCT 1778
 DB 901 TTCCCTGCAACCTCAGAGACCAAGGACCTGCTGAGTGTGTTTCCAGCATCT 958

RESULT 11
 AL562576/c 1001 bp mRNA linear EST 31-MAY-2003
 LOCUS AL562576 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DC013Y021 3-PRIME, mRNA sequence.

ACCESSION AL562576
 VERSION AL562576.2 GI:31286588
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1. (bases 1 to 1001)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 15, 2001 this sequence version replaced gi:12911134.

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 10000.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DC013CE11NP1&cluster=10000.f. Contact :
 Feng Liang Email : fliang@life.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0DC013CE11NP1.
 Location/Qualifiers

FEATURES

source

1. 1001
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC013Y021"
 /issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 40.8%; Score 862.4; DB 9; Length 1001;
 Best Local Similarity 94.0%; Pred. No. 6.3e-154;
 Matches 942; Conservative 33; Mismatches 18; Indels 9; Gaps 8;

QY 1081 AGGATCTCTCTGACACCCCTGCTGCGACAGAGACCCCATAGCCCTTAAGCCCAACCGATC 1140
 DB 1000 AAGCATCTCTCTGACACCCCTGCTGCGACAGAGACCCCATAGCCCTTAAGCCCAACCGATC 941

```

QY 1141 CCATCTCTGGAGGCTGCCAGGTGTCCTCTGATGATGACTGGGGCTGAGACGAGCCAGGGA 1200
Db 940 YVATCTTCTGGAGGCTGCCAGGTGTCCTCTGATGATGACTGGGGCTGAGACGAGCCAGGGA 881
QY 1201 AGAGAGGAGGAGACAGAGAGTGGTCTGTATGCTGATGAGACACCTCTACTACACGCTCAGC 1260
Db 880 AGAGAGGAGGAGACAGAGAGTGGTCTGTATGCTGATGAGACACCTCTACTACACGCTCAGC 821
QY 1261 TGGCAACAGTGAATTGATTTGGGGGTAGCTCCAGCCTTCTCTGCTCTGAACTGAGC 1320
Db 820 TGCACACAGTGAATTGATTTGGGGGTAGCTCCAGCCTTCTCTGCTCTGAACTGAGC 762
QY 1321 GAAACCTTCAAGTCCCTTCCAGAGGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 761 GAAACCTTCAAGTCCCTTCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702
QY 1381 ACACATCTGCTTGTGTCACACA-CATGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1439
Db 701 ACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 642
QY 1440 GCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1498
Db 641 GCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 582
QY 1499 CAGAGATGACAAACATG-GCATCTCTGAGCTGACAACTCTTTCATGACCAATAGGTC-AC 1556
Db 581 CAGAGATGACAAACATG-GCATCTCTGAGCTGACAACTCTTTCATGACCAATAGGTC-AC 522
QY 1557 TGTCTACACTGGGTACACTTTGTACCACTGTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1616
Db 521 TGTCTACACTGGGTACACTTTGTACCACTGTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 462
QY 1617 CCTCTGTCAGAGCAATCCTCTTTCACAAACAAACAGAGCTGC--TTTGATCTTGTATAC 1674
Db 461 CCTCTGTCAGAGCAATCCTCTTTCACAAACAAACAGAGCTGC--TTTGATCTTGTATAC 402
QY 1675 TTTTCAGAGAGAGAGAGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1734
Db 401 TTTTCAGAGAGAGAGAGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 342
QY 1735 GAGACCCAGAGCCAGCTCACTCTGAGAACTGTGTCCAGAGACTCTCTGCTGCTGCTGCTGCTG 1794
Db 341 GAGACCCAGAGCCAGCTCACTCTGAGAACTGTGTCCAGAGACTCTCTGCTGCTGCTGCTGCTG 282
QY 1795 GAATTTCTCTCTGAGGCTTAAGCTTGGGATTTGGGGCAGAGATTAAGATTCAGAACTATG 1854
Db 281 GAATTTCTCTCTGAGGCTTAAGCTTGGGATTTGGGGCAGAGATTAAGATTCAGAACTATG 222
QY 1855 AGGCTAGTCTTGTCTAATCACTCAAGACTGTCTGATGAGGGT-CCAGGCTGTCAACCA 1913
Db 221 AGGCTAGTCTTGTCTAATCACTCAAGACTGTCTGATGAGGGT-CCAGGCTGTCAACCA 162
QY 1914 TGGGGCTTCTGAGCTGAGCAACAGGTTGAGGAGAGAGATTAAGCAAGGCTGTCTGCTGCTG 1973
Db 161 TGGGGCTTCTGAGCTGAGCAACAGGTTGAGGAGAGAGATTAAGCAAGGCTGTCTGCTGCTG 102
QY 1974 GCCACCTGAGAAAGTCCAGGCTGAGCTCTTCTGAGGAGCACTTGGGCTGTCAATCCAGG 2033
Db 101 GCCACCTGAGAAAGTCCAGGCTGAGCTCTTCTGAGGAGCACTTGGGCTGTCSMAATCCOMMA 42
QY 2034 TCCATCTCTAGGTTTGGATATACATGATGATGATGATGATGATGATGATGATGATGATG 2075
Db 41 CCNNNTANNNTAGGTTT-CAATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1

```

```

RESULT 12
LOCUS BQ216198 873 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT_7574831 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6059905
ACCESSION BQ216198
VERSION BQ216198.1 GI:20397598
KEYWORDS EST.

```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: L14M13327 row: f column: 02
High quality sequence stop: 640.
FEATURES
source
location/Qualifiers
1..873
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6059905"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 68"
/note="Organ: lung; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."

```

```

ORIGIN
Query Match 40.2%; Score 849.8; DB 13; Length 873;
Best Local Similarity 99.1%; Pred. No. 1.6e-151;
Matches 865; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

```

```

QY 111 AGGCTCTGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 170
Db 1 AGGCTCTGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 171 CCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 230
Db 61 CCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 231 ACTTAGATACCGAGCGTCCCGTCCAGAAAGAGCTGGAAGTGGGCCCCAGCCAGACTGC 290
Db 121 ACTTAGATACCGAGCGTCCCGTCCAGAAAGAGCTGGAAGTGGGCCCCAGCCAGACTGC 180
QY 291 CCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 350
Db 181 CCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 351 CTGCGCTCCGCTTGGGGCCCTATGCTCTCTGAGAGCCCGAGAGGGCGGGCGGCTTACC 410
Db 241 CTGCGCTCCGCTTGGGGCCCTATGCTCTCTGAGAGCCCGAGAGGGCGGGCGGCTTACC 300
QY 411 AGGCTCTGAGCTCCCTACAGGCACTAGATATACCTGCAAGGTTACCCCGTCCAGGAAG 470
Db 301 AGGCTCTGAGCTCCCTACAGGCACTAGATATACCTGCAAGGTTACCCCGTCCAGGAAG 360
QY 471 CCCCTGCGCTGCTGAGAGCCCTTACGCGGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 530
Db 361 CCCCTGCGCTGCTGAGAGCCCTTACGCGGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 531 CTGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 590
Db 421 CTGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 591 TGCAAGAGCTGCTGAGAGAGCGCGCACTATCCCTGAGAGCTGAGAGCTGCGGCTCTTCC 650
Db 481 TGCAAGAGCTGCTGAGAGAGCGCGCACTATCCCTGAGAGCTGAGAGCTGCGGCTCTTCC 540

```


PE 14-SEP-2000; 2000MO-US025515.
 XX
 PR 15-SEP-1999; 99US-0154141P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Tang YT, Bandman O, Hillman JL, Baughn MF, Azimzai Y,
 PI Lu DAM;
 XX
 DX WPI; 2001-244811/25.
 DR N-PSDB; AAF30480.
 XX
 PT Novel human protein phosphatase and kinase proteins for diagnosis,
 PT treatment and prevention of gastrointestinal, immune system, neurological
 PT and cell proliferative disorders.
 XX
 XS Claim 1; Page 88-89; 103pp; English.

CC The present sequence is that of novel human protein phosphatase and
CC kinase protein PPHK-P-5, as predicted from Inyte Clone ID No. 1271505CB1
CC (see AAF030480). Tissues that express PPHK-P-5 (as a fraction of total
CC tissues expressing PPHK-P-5) include reproductive (0.288),
CC gastrointestinal (0.212) and haematopoietic or immune (0.192). Diseases
CC or conditions associated with tissues expressing PPHK-P-5 (as a fraction
CC of total tissues expressing PPHK-P-5) include cancer (0.577), inflammation
CC or trauma (0.327) and cell proliferation (0.308). The encoded protein
CC shows homology to rat kinase. The invention provides human PPHK-P-1 to -11
CC polypeptides (see AAB20322-32) and polynucleotides (see AAF030476-86). It
CC also provides expression vectors, host cells, antibodies, agonists and
CC antagonists, as well as methods for diagnosing, treating or preventing
CC disorders associated with expression of PPHK-P, including gastrointestinal
CC disorders, immune system disorders, neurological disorders and cell
CC proliferative disorders, including cancer

Sequence 358 AA;
SQ

Query Match	100.0%	Score 1892;	DB 4;	Length 358;
Best Local Similarity	100.0%	Pred. No. 6.1e-176;		
Matches 358;	Conservative 0;	Mismatches 0;	Gaps 0;	

QY	MRATPLAAPAGSLSRKKRLIEDDNLDTEPVOKARSGORPLRCLPLSPFPAPDRAT	60
Db	1 MRATPLAAPAGSLSRKKRLIEDDNLDTEPVOKARSGORPLRCLPLSPFPAPDRAT	60
QY	61 AVATASRLGPVLLLEPEGSGRAYVALHCPGTGYTCCKYPQVALAVLEPYARLPKHKV	120
Db	61 AVATASRLGPVLLLEPEGSGRAYVALHCPGTGYTCCKYPQVALAVLEPYARLPKHKV	120
QY	121 ARTEVLAGTGLLYAFETTHGDMISIVRTRIRIPREPAVLTFQMATAALAHCHQHLVL	180
Db	121 ARTEVLAGTGLLYAFETTHGDMISIVRTRIRIPREPAVLTFQMATAALAHCHQHLVL	180
QY	181 RDIKLCRPVADERRKKVLENILEDSCVLTGDDSLMDKHACPAYVGEILSSRASYSKG	240
Db	181 RDIKLCRPVADERRKKVLENILEDSCVLTGDDSLMDKHACPAYVGEILSSRASYSKG	240
QY	241 AADVMSLGAALFTMLAGHYFPDQSEPVLLFGKIRRGAVLAPAGLSAPAPCLVRCLLREP	3000
Db	241 AADVMSLGAALFTMLAGHYFPDQSEPVLLFGKIRRGAVLAPAGLSAPAPCLVRCLLREP	3000
QY	301 AERLTATGTLHPWLRQDPMPLAFTRSHLMEAAQVPPDGLIDAREEBSGREVLYVG	358
Db	301 AERLTATGTLHPWLRQDPMPLAFTRSHLMEAAQVPPDGLIDAREEBSGREVLYVG	358

RESULT 2	
AAU03509	
XX	AAU03509 standard; protein; 358 AA.
XX	
AC	AAU03509;
XX	
DT	12-SEP-2001 (first entry)
XX	

DE Human protein kinase #9.
XX
KW Human; protein kinase; PK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.

OS	Homo sapiens.
XX	
PN	WO200138503-A2.
XX	
PD	31-MAY-2001.

PF 22-NOV-2000; 2000WO-US032085.

PR 24-NOV-1999; 99US-0167482P.

PA (SUGE-) SUGEN INC.

Plowman GD, Whyte D,

100

DR N-PSDB; AAS06709.

PT Nucleic acids encoding human kinase polypeptides, useful for preventing

PT neuronal-associated diseases, and microbial infections.

PS Claim 7; Fig 2; 433pp; English.

AA00350, AA00357 represent novel human protein kinases #1-57. The novel protein kinases have been identified as members of the tyrosine or serine/threonine kinase (PTK and STK) families. The polynucleotides encoding protein kinases and the polypeptides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate kinase expression. For example, they may be used to treat cancers (especially cancers of haematopoietic origin), cardiovascular disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes), immune related diseases (e.g. rheumatoid arthritis), neurological disorders (e.g. schizophrenia), neurodegenerative disorders (e.g. Parkinson's disease), inflammatory disorders (e.g. asthma), infectious disease (e.g. HIV) and reproductive disorders (e.g. infertility). Additionally, polynucleotides encoding protein kinases may be used for gene therapy and as DNA probes in diagnostic assays. The protein kinase polypeptides may be used as antigens in the production of antibodies against the protein kinases and in assays to identify modulators of protein kinase expression and activity

Sequence 358 AA;
SQ

Query Match	99.8%	Score 1888	DB 4	Length 358
Best Local Similarity	99.7%	Pred. NO. 1.5e-175		
Matches 357	Conservative 1	Mismatches 0	Indels 0	Gaps 0

QY	1	MAATPLAAAGSLSKRRKLELDNDLPTERRPOKARSGSPORLPECLLESPFPAPRAT	60
QY	1	MATPLPAAAGSLSKRRKLELDNDLPTERRPOKARSGSPORLPECLLESPFPAPRAT	60
Db	61	AAVATASRLGPVYLLEPEEGSGRAYQALHCFGTETVTCCKYVPOEALAVLEPYARLPEPKHV	120
QY	61	AAVATASRLGPVYLLEPEEGSGRAYQALHCFGTETVTCCKYVPOEALAVLEPYARLPEPKHV	120
Db	61	AAVATASRLGPVYLLEPEEGSGRAYQALHCFGTETVTCCKYVPOEALAVLEPYARLPEPKHV	120
QY	121	AAPETEVLAGTQLLYAFRTRTGDMDSLYTRSHRIPEPEAAVLFPQOMTALAHCHQHSVL	180
QY	121	AAPETEVLAGTQLLYAFRTRTGDMDSLYTRSHRIPEPEAAVLFPQOMTALAHCHQHSVL	180
Db	121	AAPETEVLAGTQLLYAFRTRTGDMDSLYTRSHRIPEPEAAVLFPQOMTALAHCHQHSVL	180
QY	181	RLULKCRFPFADREKKKVLLENLEBSCVLTGPDDSLMDKACPAVYGGEBILSSPAYSYGK	240
QY	181	RLULKCRFPFADREKKKVLLENLEBSCVLTGPDDSLMDKACPAVYGGEBILSSPAYSYGK	240
Db	181	RLULKCRFPFADREKKKVLLENLEBSCVLTGPDDSLMDKACPAVYGGEBILSSPAYSYGK	240
QY	241	AAADVSLGVALFTMLAGHYPFQDSSEPVLLFGKIRGAYVALPAGIASAPARCIVRCLLREEP	300
QY	241	AAADVSLGVALFTMLAGHYPFQDSSEPVLLFGKIRGAYVALPAGIASAPARCIVRCLLREEP	300

Db	241	AADWWSLGVAFETMLAGHPFQDSEPVLLFGKIRRGAVYLLPGLSAPACIVRCILRRBP	301
Qy	301	AERLTATGILLHPMLRQDPMPIAPFRSHIMEAAQVVPDGLGIDEAREEGDREVVL	358
Db	301	AERLTATGILLHPMLRQDPMPIAPFRSHIMEAAQVVPDGLGIDEAREEGDREVVL	358
RESULT 3			
AA38908			
AA38908		standard; protein; 358 AA.	
AA38908;			
22-OCT-2001		(first entry)	
Human polypeptide SEQ ID NO 2053.			
Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.			
Homo sapiens.			
WO200153312-A1.			
26-JUL-2001.			
26-DEC-2000; 2000MO-US034263.			
23-DEC-1999; 99US-00471275.			
21-JAN-2000; 2000US-00488725.			
25-APR-2000; 2000US-00552317.			
20-JUN-2000; 2000US-00598042.			
19-JUL-2000; 2000US-00620312.			
03-AUG-2000; 2000US-00653450.			
14-SEP-2000; 2000US-00662191.			
19-OCT-2000; 2000US-00693036.			
29-NOV-2000; 2000US-00727344.			
(HYBE-) HYSEQ INC.			
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D, Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA, Zhou F, Goodrich R, Drmanac R;			
WPI; 2001-442253/47.			
N-PSDB; AA158064.			
Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.			
Example 3; SEQ ID NO 2053; 10078pp; English.			
The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AA38642-AA42213) with noctropic.			
immunosuppressant and cyrostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification			
Sequence 358 AA;			

Query Match	99.6%;	Score 1885;	DB 4;	Length 358;
Best Local Similarity	99.4%;	Pred. No. 2.9e-175;		
Matches 356;	Conservative 2;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MRATPLAAPAGSISRRKRLIEDNLDTERVPYQKRAASGPORPLRPCLLPSPPTAPDRAT	60	
Db	1	MRATPLAAPAGSISRRKRLIEDNLDTERVPYQKRAASGPORPLRPCLLPSPPTAPDRAT	60	
QY	61	AVATASLGRVYLIEPEEGRAAYALHCPGTETVCTCKVYVQGLAVLEPYARLPKHV	120	
Db	61	AVATASLGRVYLIEPEEGRAAYALHCPGTETVCTCKVYVQGLAVLEPYARLPKHV	120	
QY	121	ARPEVLVAGQQLYAFPTTHGDMHSVTRFRHRIPEEAAVLFPROMATALAHCHQGLVL	180	
Db	121	ARPEVLVAGQQLYAFPTTHGDMHSVTRFRHRIPEEAAVLFPROMATALAHCHQGLVL	180	
QY	181	RDCLKCFVPADREKKLVLENIEDSCVLTGPDDSLMDKACPAVYGEELLSSRAVSXGK	240	
Db	181	RDCLKCFVPADREKKLVLENIEDSCVLTGPDDSLMDKACPAVYGEELLSSRAVSXGK	240	
QY	241	AADVMSIGVALFTMLAGHYPRODESPVILPEKIRGAYALPAGSAPARCLVRCILRREP	300	
Db	241	AADVMSIGVALFTMLAGHYPRODESPVILPEKIRGAYALPAGSAPARCLVRCILRREP	300	
QY	301	AERLTATGILILHPWLRODPMELAPTRSHLWEAAQVDPDGLIDEAREEESGDRVVLYG	358	
Db	301	AERLTATGILILHPWLRODPMELAPTRSHLWEAAQVDPDGLIDEAREEESGDRVVLYG	358	
RESULT 4				
AA040694				
ID	AA040694	standard; protein; 393 AA.		
AC	AA040694;			
XX				
DT	22-OCT-2001	(first entry)		
DE		Human polypeptide SEQ ID NO 5625.		
XX				
KM		Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KM		peripheral nervous system; neuropathy; central nervous system; CNS;		
KM		Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KM		amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemolactic;		
KM		chemokineic; thrombolytic; drug screening; arthritis; inflammation;		
KM		leukemia.		
XX				
OS		Homo sapiens.		
XX				
PN	WO200153312-A1.			
XX				
PD	26-JUL-2001.			
XX				
PF	26-DEC-2000;	2000WO-US034263.		
XX				
PR	23-DEC-1999;	99US-00471275.		
PR	21-JAN-2000;	2000US-00488725.		
PR	25-APR-2000;	2000US-0052317.		
PR	20-JUN-2000;	2000US-00589042.		
PR	19-JUL-2000;	2000US-00620312.		
PR	03-AUG-2000;	2000US-00653450.		
PR	14-SEP-2000;	2000US-00662191.		
PR	19-OCT-2000;	2000US-00693036.		
PR	29-NOV-2000;	2000US-00727344.		
XX				
PA	(HYSE-) HYSEQ INC.			
XX				
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;			
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Q;			
PI	Zhou P, Goodrich R, Dymnac RT;			
XX				
DR	WPI: 2001-442253/47.			
DR	N-PSDB; AA159650.			

XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Example 2; SEQ ID NO 5625; 1007bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AA158642-AA164213) with neurotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemoclastic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX

Sequence 393 AA;

Query Match 99.4%; Score 1885; DB 4; Length 393;
 Best Local Similarity 99.4%; Pred. No. 3.3e-175;
 Matches 356; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSRKKRLDNDLTERPVQKRAAGQPPRLPCLPLSPPTADPRAT 60
 DB 36 MRATPLAAPAGSLSRKKRLDNDLTERPVQKRAAGQPPRLPCLPLSPPTADPRAT 95
 QY 61 AVATASRLGPGVYLLEPEEGGAYOALHCPGTGYTCVVVQVQALAVLEPYARLPKHKV 120
 DB 96 AVATASRLGPGVYLLEPEEGGAYOALHCPGTGYTCVVVQVQALAVLEPYARLPKHKV 155
 QY 121 ARPEVLAGTQLLYAFPTRTGDMHSLVTRRRIPEPEAAVLFRQATLALHCHQGLVL 180
 DB 156 ARPEVLAGTQLLYAFPTRTGDMHSLVTRRRIPEPEAAVLFRQATLALHCHQGLVL 215
 QY 181 RDLKICRFEPADREKKVLLENEDSCVLTGPDSDLMQKACAYVGPPEILSSRASYSKG 240
 DB 216 RDLKICRFEPADREKKVLLENEDSCVLTGPDSDLMQKACAYVGPPEILSSRASYSKG 275
 QY 241 AADVMSLGVALTMLAGHYFPQDSPEVLLFGKIRRGAVYALPAGLSAPACLVRCILRREP 300
 DB 276 AADVMSLGVALTMLAGHYFPQDSPEVLLFGKIRRGAVYALPAGLSAPACLVRCILRREP 335
 QY 301 AERLTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVPDGLGDEAREEGDEEVLYG 358
 DB 336 AERLTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVPDGLGDEAREEGDEEVLYG 393

RESULT 5
 AACS9337
 ID AACS9337 standard; protein; 358 AA.
 AC AACS9337;
 XX

DT 18-DEC-2003 (first entry)

XX Human cell death inhibitory protein.

XX Human; cell-death; cardiatic; neuroprotective; anti-HIV; antiinflammatory;
 DE cerebroprotective; hepatotropic; cytostatic; immunosuppressive;
 XX anti-rheumatic; ophthalmological; neurotropic; antiparkinsonian;
 XX anticonvulsant; hypotensive; antiarteriosclerotic; haemostatic;
 XX antialcoholic; virocidic; HIV; cardiac disease; immunological disease;
 XX neurodegenerative disease; ischaemic damage;
 XX central nervous system disorder; hepatic disorder;
 XX acute pancreatic inflammation; cancer; AIDS; autoimmune disease;
 XX rheumatism; Crohn's disease; glaucoma; Alzheimer's disease;
 XX Parkinson's disease; Huntington's disease; hypertension;

KW arteriosclerosis; reperfusion damage; myocardial infarction;
 KW cerebral trauma; cerebral infarction; cerebral haemorrhage; hepatitis;
 KW alcoholic hepatitis; cerebral ischemia.
 XX
 XX Homo sapiens.

XX JP2003063986-A.

XX 05-MAR-2003.

XX 27-AUG-2001; 2001JP-00255811.

XX 27-AUG-2001; 2001JP-00255811.

XX (SUMT) SUMITOMO SEIYAKU KK.

XX WPI; 2003-601360/57.

XX N-PDB; AACS9338.

PT Controlling cell-death by administering positive or negative regulator of
 PT cell-death inhibitory-factor, for treating cancer, AIDS, autoimmune
 PT diseases, Crohn's diseases, glaucoma, Alzheimer disease.

PS Claim 1; SEQ ID NO 3; 35pp; Japanese.

CC This invention relates to a novel method for controlling cell-death
 CC comprising administering a positive or negative regulator of cell-death
 CC inhibitory-factor. The invention also comprises a method for screening
 CC modulators of cell-death, by contacting cells expressing cell-death
 CC inhibitory factor with a candidate compound, monitoring level of
 CC expression of cell-death inhibitory factor, evaluating cell-death
 CC modulation ability of the compound based on change in the level of
 CC expression of the factor and selecting compounds having cell-death
 CC modulation ability. The cell death regulator of the invention may have
 CC cardiatic, neuroprotective, anti-HIV, antiinflammatory, cerebroprotective,
 CC hepatotropic, cytostatic, immunosuppressive, antiparkinsonia, anticonvulsant,
 CC ophthalmological, neurotropic, antiparkinsonia, antineumatic,
 CC hypotensive, antiarteriosclerotic haemostatic, antialcoholic and
 CC virocidic activities. The method of the invention is useful for treating
 CC HIV, cardiac diseases, immunological diseases, neurodegenerative disease,
 CC ischaemic damage and congestion, disorder of central nervous system,
 CC hepatic disorder, acute pancreatic inflammation, and cancer, AIDS,
 CC autoimmune diseases, rheumatism, Crohn's disease, glaucoma, Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, hypertension,
 CC arteriosclerosis, reperfusion damage, myocardial infarction, cerebral
 CC trauma, cerebral infarction, cerebral haemorrhage, hepatitis, alcoholic
 CC hepatitis, and cerebral ischemia. The present sequence represents the
 CC human cell death inhibitory protein of the invention.
 XX

Sequence 358 AA;

Query Match 99.4%; Score 1881; DB 7; Length 358;
 Best Local Similarity 99.4%; Pred. No. 7.2e-175;
 Matches 356; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSRKKRLDNDLTERPVQKRAAGQPPRLPCLPLSPPTADPRAT 60
 DB 1 MRATPLAAPAGSLSRKKRLDNDLTERPVQKRAAGQPPRLPCLPLSPPTADPRAT 60
 QY 61 AVATASRLGPGVYLLEPEEGGAYOALHCPGTGYTCVVVQVQALAVLEPYARLPKHKV 120
 DB 61 AVATASRLGPGVYLLEPEEGGAYOALHCPGTGYTCVVVQVQALAVLEPYARLPKHKV 120
 QY 121 ARPEVLAGTQLLYAFPTRTGDMHSLVTRRRIPEPEAAVLFRQATLALHCHQGLVL 180
 DB 121 ARPEVLAGTQLLYAFPTRTGDMHSLVTRRRIPEPEAAVLFRQATLALHCHQGLVL 180
 QY 181 RDLKICRFEPADREKKVLLENEDSCVLTGPDSDLMQKACAYVGPPEILSSRASYSKG 240
 DB 181 RDLKICRFEPADREKKVLLENEDSCVLTGPDSDLMQKACAYVGPPEILSSRASYSKG 240
 QY 241 AADVMSLGVALTMLAGHYFPQDSPEVLLFGKIRRGAVYALPAGLSAPACLVRCILRREP 300
 DB 241 AADVMSLGVALTMLAGHYFPQDSPEVLLFGKIRRGAVYALPAGLSAPACLVRCILRREP 300

DB 241 AADVWSLGVALLFTMLAGHYPPFODSEPVLLFGKIRRGAYALPAGLSAPARCLVRCILRRP 300
QY 301 AERLTATGILLHPMLRQDPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358
DB 301 AERLTATGILLHPMLRQDPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358
RESULT 6
ID ABB80976 standard; protein; 360 AA.
AC ABB80976;
XX
XX
DT 21-OCT-2002 (first entry)
DE Human tribbles homologue-3 (htrb-3) polypeptide encoding DNA.
XX
XX Human; tribbles; htrb-1; stress kinase inhibitor protein; SKIP-1; AP-1;
KW anti-rheumatic; antiarthritic; antidiabetic; antipsoriatic; osteopathic;
KW ophthalmological; cardiac; cytosolic; haemostatic; immunosuppressive;
KW anti-inflammatory; estrogen receptor; fibroblast growth factor; FGF;
KW tumor necrosis factor; TNF; htrb-3.
XX
XX Homo sapiens.
OS
PN WO200253743-A2.
XX
XX 11-JUL-2002.
PD
PF 08-JAN-2002; 2002WO-US000070.
XX
PR 08-JAN-2001; 2001US-0260294P.
XX
PA (INTE-) INTERLEUKIN GENETICS INC.
PI Dower S, Qunastrom E, Kiss-Toth E;
XX WPI; 2002-590635/63.
DR N-PSDB; ABB86479.
XX
XX Novel isolated human tribbles homologue-1 polypeptide for inhibiting AP-1
PT -mediated inflammatory signal in a cell, and activating ERK-mediated
PT signal e.g. AP-1-mediated gene activation signal in a cell.
XX
XX Example 10; Fig 11B; 131pp; English.
PS
XX The invention provides an isolated human tribbles homologue-1 (htrb-1,
CC also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
CC htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory
CC signal in a cell. The polypeptide employed in the method is preferably
CC htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N htrb-3 C, or htrb
CC -3 N C. It is also useful for providing htrb agonist activity for
CC activating an ERK-mediated signal e.g. AP-1-mediated gene activation
CC signal, an estrogen receptor-mediated gene activation signal, an
CC fibroblast growth factor (FGF) induced signal, or a PMA induced signal,
CC in a cell. Htrb modulators are useful for modulating AP-1 mediated
CC inflammatory signal in a cell such as tumor necrosis factor (TNF) induced
CC inflammatory signal, or an interleukin induced inflammatory signal. htrb
CC proteins are useful in screening assays, predictive medicine and in
CC compounds e.g. for treating and/or preventing diseases caused by abnormal
CC htrb activity, such as rheumatoid arthritis, diabetes, psoriasis,
CC osteoporosis, diabetic retinopathy, myocardial infarction and cancers.
CC The htrb therapeutics are useful for antagonizing interleukin-1 dependent
CC disorders of human placenta, intraventricular hemorrhage, neonatal white
CC matter damage and subsequent cerebral palsy, and inflammation or
CC autoimmune disorders. The present sequence represents the htrb-3
CC polypeptide
XX
SQ Sequence 360 AA;
Query Match 98.9%; Score 1871; DB 5; Length 360;
Best Local Similarity 98.6%; Pred. No. 6.9e-1/4;

Matches 355; Conservative 3; Mismatches 0; Indels 2; Gaps 1;
QY 1 MRATPLAAPAGSLSRKKRLIEDNDLTERPVQKARSGPQRLPLSPPTAPDRA 60
DB 1 MRATPLAAPAGSLSRKKRLIEDNDLTERPVQKARSGPQRLPLSPPTAPDRA 60
QY 61 AVATASRLGPVLLPEPEGGAAYQALHCPCTCTCKYYPQVQALAVLEPARLPKHGV 120
DB 61 AVATASRLGPVLLPEPEGGAAYQALHCPCTCTCKYYPQVQALAVLEPARLPKHGV 120
QY 121 ARPTVLGTOQLVAFPTRTGDMSLVTRHRIPEPEAAVLEFQMATALACHOHGVL 180
DB 121 ARPTVLGTOQLVAFPTRTGDMSLVTRHRIPEPEAAVLEFQMATALACHOHGVL 180
QY 181 RDLKLCREVEFA--DREKKVLLENLEDSCVLTGPDSDLMKACPAVYGPETLSSRASYS 238
DB 181 RDLKLCREVEFA--DREKKVLLENLEDSCVLTGPDSDLMKACPAVYGPETLSSRASYS 238
QY 239 GKADVWSLGVALLFTMLAGHYPPFODSEPVLLFGKIRRGAYALPAGLSAPARCLVRCILRR 298
DB 241 GKADVWSLGVALLFTMLAGHYPPFODSEPVLLFGKIRRGAYALPAGLSAPARCLVRCILRR 300
QY 299 BPAERLTATGILLHPMLRQDPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358
DB 301 BPAERLTATGILLHPMLRQDPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 360
RESULT 7
ID AAY69157 standard; peptide; 360 AA.
XX
XX AAY69157;
XX
XX 30-MAY-2000 (first entry)
DT
DE Peptidea J5503-KS comprising domains VIA to XI of a protein kinase.
XX
XX Peptide activity; molecular weight marker; isoelectric focusing marker;
KW kinase peptide fragmentation control; cellular signal transduction.
XX
XX Homo sapiens.
OS
XX WO200008180-A2.
XX
XX 17-FEB-2000.
PD
PF 03-AUG-1999; 99WO-US017630.
XX
XX 04-AUG-1998; 98US-0095270P.
PR 11-SEP-1998; 98US-0099972P.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Virca GD, Bird TA, Anderson DM, Marken JS;
XX WPI; 2000-195584/17.
DR N-PSDB; AAY61155.
XX
XX New human kinase polypeptides and polynucleotides used as molecular
PT weight markers and as controls for peptide fragmentation.
XX
PS Claim 2; Page 10; 60pp; English.
XX
XX The present sequence represents a partial polypeptide which has kinase
CC activity. The kinase polynucleotides can be used to express the
CC polypeptides, and as probes to identify nucleic acids encoding proteins
CC having kinase activity. The kinase polypeptides and fragmented
CC polypeptides are used as molecular weight and isoelectric focusing
CC markers, and as controls for peptide fragmentation. They also have a
CC number of therapeutic uses as kinases play a central role in cellular
CC signal transduction. The polypeptides could also be used to identify
CC binding partner proteins. The polypeptides can also be used as a reagent
CC to identify any proteins that the polypeptide regulates, and proteins

CC with which it might interact. The polypeptides may also be used for
 CC preparation of antibodies. The antibodies can be used in assays to detect
 CC the presence of the protein, and to purify the protein by immunofluorescence
 CC chromatography

XX Sequence 360 AA:

Query Match 88.6%; Score 1676; DB 3; Length 360;
 Best Local Similarity 99.7%; Pred. No. 7,8e-155;
 Matches 317; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSRKKRLLELDNLDTERPVQKRAAGPQRLPCLLPSPPTAPDRAT 60
 DB 43 MRATPLAAPAGSLSRKKRLLELDNLDTERPVQKRAAGPQRLPCLLPSPPTAPDRAT 102
 QY 61 AVATASRLGPGYVLEPEEGRAYQALHCPGTGYTCVYVQALAVLEBYARLPKHV 120
 DB 103 AVATASRLGPGYVLEPEEGRAYQALHCPGTGYTCVYVQALAVLEBYARLPKHV 162
 QY 121 ARPTVLATGTLVAFPTTRTHGDMHSLVTRHRIPEPEAAVLFROMATLAHQHGLVL 180
 DB 163 ARPTVLATGTLVAFPTTRTHGDMHSLVTRHRIPEPEAAVLFROMATLAHQHGLVL 222
 QY 181 RDLKLCRFVADREKKLVLENLEDSCVLTGPDSDLMDKACPAVYGPETLSSRASYSK 240
 DB 223 RDLKLCRFVADREKKLVLENLEDSCVLTGPDSDLMDKACPAVYGPETLSSRASYSK 282
 QY 241 AADVMSLGVALLFTMLAGHYFQDSEPVLLFGKIRGAYALPAGLSAPARCLVCLRRP 300
 DB 283 AADVMSLGVALLFTMLAGHYFQDSEPVLLFGKIRGAYALPAGLSAPARCLVCLRRP 342
 QY 301 AERLTATGILLHPWLROD 318
 DB 343 AERLTATGILLHPWLROD 360

RESULT 8
 AAB85791
 ID AAB85791 standard; protein; 323 AA.
 AC AAB85791;
 DT 29-OCT-2001 (first entry)
 DE Human kinase PKIN-10.
 DE Human kinase PKIN-10.
 KM PKIN; kinase; cytosolic; immunosuppressive; immunostimulant; human;
 KM antiarteriosclerotic; cardiact; gene therapy; antisense therapy.
 OS Homo sapiens.
 OS Homo sapiens.
 OS WO200160991-A2.
 PN 23-AUG-2001.
 PD 16-FEB-2001; 2001MO-US005240.
 PF 17-FEB-2000; 2000US-0183682P.
 PR 02-MAR-2000; 2000US-0186559P.
 PR 09-MAR-2000; 2000US-0188606P.
 PR 17-MAR-2000; 2000US-0189989P.
 PR 30-MAR-2000; 2000US-0193851P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Tang YT, Buford N, Gandhi AR, Patterson C, Khan JA, Yue H;
 PI Hafalia A, Shih LT, Tribouley CM, Yao MG, Burdill JD, Marcus GA;
 PI Zingler KA, Lu DAM, Bandman O, Policky JL, Griffin JA, Thornton M;
 PI Nguyen DB, Lai P, Walsh RT;
 XX WPI: 2001-514771/56.
 DR N-PSDB; AAB76218.
 XX

PT isolated human kinase polypeptides useful in the diagnosis, treatment and
 PT prevention of cancer, immune disorders and disorders affecting growth and
 PT development.

XX Claim 1; Page 115; 126pp; English.

CC The invention provides human kinases (PKIN) and polynucleotides encoding
 CC PKIN. The PKIN polypeptides can be expressed using standard recombinant
 CC methodology. The PKIN polypeptides, polynucleotides, modulators and
 CC specific antibodies are useful in the diagnosis, treatment and prevention
 CC of cancer, immune disorders, disorders affecting growth and development,
 CC atherosclerosis, and other cardiovascular diseases, and lipid disorders
 CC and in the assessment of the effects of exogenous compounds on the
 CC expression of nucleic acid sequences of human kinases. The present
 CC sequence represents a human PKIN-10 polypeptide

XX Sequence 323 AA:

Query Match 83.8%; Score 1585.5; DB 4; Length 323;
 Best Local Similarity 86.3%; Pred. No. 4.7e-146;
 Matches 309; Conservative 3; Mismatches 11; Indels 35; Gaps 2;

QY 1 MRATPLAAPAGSLSRKKRLLELDNLDTERPVQKRAAGPQRLPCLLPSPPTAPDRAT 60
 DB 1 MRATPLAAPAGSLSRKKRLLELDNLDTERPVQKRAAGPQRLPCLLPSPPTAPDRAT 60
 QY 61 AVATASRLGPGYVLEPEEGRAYQALHCPGTGYTCVYVQALAVLEBYARLPKHV 120
 DB 61 AVATASRLGPGYVLEPEEGRAYQALHCPGTGYTCVYVQALAVLEBYARLPKHV 120
 QY 121 ARPTVLATGTLVAFPTTRTHGDMHSLVTRHRIPEPEAAVLFROMATLAHQHGLVL 180
 DB 121 ARPTVLATGTLVAFPTTRTHGDMHSLVTRHRIPEPEAAVLFROMATLAHQHGLVL 180
 QY 181 RDLKLCRFVADREKKLVLENLEDSCVLTGPDSDLMDKACPAVYGPETLSSRASYSK 240
 DB 157 RDLKLCRFVADREKKLVLENLEDSCVLTGPDSDLMDKACPAVYGPETLSSRASYSK 205
 QY 241 AADVMSLGVALLFTMLAGHYFQDSEPVLLFGKIRGAYALPAGLSAPARCLVCLRRP 300
 DB 206 AADVMSLGVALLFTMLAGHYFQDSEPVLLFGKIRGAYALPAGLSAPARCLVCLRRP 265
 QY 301 AERLTATGILLHPWLRODMPPLAPTRSHLMEAAQVVDGLGDEAREEGBDRVLYG 358
 DB 266 AERLTATGILLHPWLRODMPPLAPTRSHLMEAAQVVDGLGDEAREEGBDRVLYG 323

RESULT 9

ADCS9335
 ID ADCS9335 standard; protein; 349 AA.

AC ADCS9335;
 DT 18-DEC-2003 (first entry)

DE Rat cell death inhibitory protein.

DE Rat cell death inhibitory protein.

KM Rat; cell-death; cardiact; neuroprotective; anti-HIV; antiinflammatory;
 KM cerebroprotective; hepatotropic; cytosolic; immunosuppressive;
 KM antirheumatic; ophthalmological; nootropic; autophagosome;
 KM anticonvulsant; hypotensive; antiarteriosclerotic; haemostatic;
 KM antialcoholic; virocidic; HIV; cardiac disease; immunological disease;
 KM neurodegenerative disease; ischaemic damage;
 KM central nervous system disorder; hepatic disorder;
 KM acute pancreatic inflammation; cancer; AIDS; autoimmune disease;
 KM rheumatism; Crohn's disease; glaucoma; Alzheimer's disease;
 KM Parkinson's disease; Huntington's disease; hypertension;
 KM arteriosclerosis; reperfusion damage; myocardial infarction;
 KM cerebral trauma; cerebral infarction; cerebral haemorrhage; hepatitis;
 KM alcoholic hepatitis; cerebral ischemia.

XX Rattus sp.
 XX
 OS
 XX

PN JP2003063986-A.
 XX
 XX 05-MAR-2003.
 PD
 XX
 XX 27-AUG-2001; 2001JP-00255811.
 PF
 XX 27-AUG-2001; 2001JP-00255811.
 PR
 XX 27-AUG-2001; 2001JP-00255811.
 PA (SUMU) SUMITOMO SEIYAKU KK.
 XX
 XX WPI: 2003-601360/57.
 DR N-PSDB; ADCS9336.
 XX
 XX Controlling cell-death by administering positive or negative regulator of
 PT cell-death inhibitory-factor, for treating cancer, AIDS, autoimmune
 PT diseases, Crohn's diseases, glaucoma, Alzheimer disease.
 PT
 PS Claim 1; SEQ ID NO 1; 35pp; Japanese.
 PS
 XX This invention relates to a novel method for controlling cell-death
 CC comprising administering a positive or negative regulator of cell-death
 CC inhibitory-factor. The invention also comprises a method for screening
 CC modulators of cell-death, by contacting cells expressing cell-death
 CC inhibitory factor with a candidate compound, monitoring level of
 CC expression of cell-death inhibitory factor, evaluating cell-death
 CC modulation ability of the compound based on change in the level of
 CC expression of the factor and selecting compounds having cell-death
 CC modulation ability. The cell death regulator of the invention may have
 CC cardiant, neuroprotective, anti-HIV, antiinflammatory, cerebroprotective,
 CC hepatotropic, cyrostatic, immunosuppressive, antirheumatic,
 CC ophthalmological, nootropic, antiparkinsonia, anticonvulsant,
 CC hypotensive, antiarteriosclerotic, haemostatic, antialcohollic and
 CC vincine activities. The method of the invention is useful for treating
 CC HIV, cardiac diseases, immunological diseases, neurodegenerative disease,
 CC ischemic damage and congestion, disorder of central nervous system,
 CC hepatic disorder, acute pancreatic inflammation, and cancer, AIDS,
 CC autoimmune diseases, rheumatism, Crohn's diseases, glaucoma, Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, hypertension,
 CC arteriosclerosis, reperfusion damage, myocardial infarction, cerebral
 CC trauma, cerebral infarction, cerebral haemorrhage, hepatitis, alcoholic
 CC hepatitis, and cerebral ischemia. The present sequence represents the rat
 CC cell death inhibitory protein of the invention.
 CC
 SQ Sequence 349 AA:
 Query Match 70.0%; Score 1324.5; DB 7; Length 349;
 Best Local Similarity 72.9%; Pred. No. 1.6e-120;
 Matches 261; Conservative 28; Mismatches 60; Indels 9; Gaps 3;
 QY 1 MRATPIAAPAGSLSRKKRLLELDNLTTERPVQRRASGQPRLLPCLPLSPPTAPDRAT 60
 Db 1 MRATSIASADVPCRKKPLFDNIDECVLRVREPERGPTPL-----PPASDLSP 55
 QY 61 AVATASRLSGPVLLPEBEGRAYOALHCPGTGTYCTKVTVQNALVLEYAALPHKRY 120
 Db 56 AVAPARIRLGYILLREQGNCTYRALHCPGTGTYCTKVTVQASAOALAYARLPHQHV 115
 QY 121 ARTEVLVAGTOLLVAFPTRTGHMSLVTRHRIPEBEAVLFRQMATALAHCHQGLVL 180
 Db 116 ARPEVLVGSQLLYFTKTHGDSHVSRRGIPPEBAALRQMSAAVAKCHKQGLL 175
 QY 181 RDLKLCFVAFADREKKLVLENEDSCVLTPGDDSLMDKACAPAYVPEILSSRASYSGK 240
 Db 176 RDLKLRFRVSNCRKTLVLENEDACVVMGPDDSLMDKACAPAYVPEILSSRPSYSGR 235
 QY 241 AADVMSIGVALFTMLAGHYPPDSEPVLLFGKTRRGAYALPAGLSAPARCLVCLRRREP 300
 Db 236 AADVMSIGVALFTMLAGRYPPDSEPALTFGKIRRGFTALPBGISAPARCLIRCLRRREP 295
 QY 301 ABRLLTAGTILLHPMLRQDMPVLAFTRSHLWEAAQVVDGGLDAREEBSGDRVLYNG 358
 Db 296 SERLVALGILLHPWLREDCGVSPPRSDREMDQVVDGGLQLEA--EEG--EVGLYNG 349

RESULT 10
 AAB43657
 ID AAB43657 standard; protein, 233 AA.
 XX
 XX AAB43657;
 AC
 XX 08-FEB-2001 (first entry)
 DT
 XX
 XX Human cancer associated protein sequence SEQ ID NO:1102.
 DE
 XX
 XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cyrostatic; proliferative; vulnerary; immunomodulator;
 KW antidiabetic; antiaschmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antihypertic; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX
 OS Homo sapiens.
 XX
 XX WO200055350-A1.
 PN
 XX 21-SEP-2000.
 PD
 XX 08-MAR-2000; 2000WO-US005882.
 PF
 XX 12-MAR-1999; 99US-0124270P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 PI WPI, 2000-587533/55.
 XX N-PSDB; AAC7866.
 DR
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer.
 PT
 PS Claim 11; Page 1707-1708; 2352pp; English.
 PS
 XX AAC78607 to AAC78448 encode the human cancer associated proteins given in
 CC AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cyrostatic; proliferative; vulnerary; immunomodulator;
 CC antidiabetic; antiaschmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antihypertic; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polypeptides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleocides, antibodies
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 CC
 SQ Sequence 233 AA:
 Query Match 58.2%; Score 1102; DB 3; Length 233;
 Best Local Similarity 100.0%; Pred. No. 5.1e-99;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 151 RHRIPPEAVALFRQMATALAHCHQGLVLRDLKLCFVAFADREKKLVLENEDSCVL 210

Db 26 RHRIPBEAAVLFRQMTALAHCHQHLVLDLKLCPVFAERKRLVLENDSCVLT 85
 QY 211 GPDSLMDKACPCPYVGEIISRSASVSGKADWMSGVAFMTLAGHYPPQDEPVLTF 270
 Db 86 GPDSLMDKACPCPYVGEIISRSASVSGKADWMSGVAFMTLAGHYPPQDEPVLTF 145
 QY 271 GKIRGAVAPAGLSAPARCLVRCILRRPAPRLTATGILHPMLRQDPMPLAFTRSHLW 330
 Db 146 GKIRGAVAPAGLSAPARCLVRCILRRPAPRLTATGILHPMLRQDPMPLAFTRSHLW 205
 QY 331 EAAQVVPDGLGLDEAREEGREYVLTG 358
 Db 206 EAAQVVPDGLGLDEAREEGREYVLTG 233

RESULT 11
 ABB06093
 ID ABB06093 standard; protein; 278 AA.
 AC ABB06093;
 DT 10-MAY-2002 (first entry)
 DE Human NS protein sequence SEQ ID NO:185.
 XX
 KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;
 KW antitubercular; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
 KW vasoregic; antiarteriosclerotic; antiinflammatory; dermatological;
 KW anorectic; muscular; antinfertilit; cardiovascular; anticonvulsant;
 KW anticonvulsant; antidiabetic; tranquiliser; immunomodulator; cardiant;
 KW gastrointestinal; virucide; antitumor; antidepressant; neuroleptic;
 KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
 KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
 KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KW infertility; cardiovascular disease; coagulation disease; hypertension;
 KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
 KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KW gastric ulcer; Alzheimer's disease.

OS Homo sapiens.
 XX
 PN WO200206315-A2.
 XX
 PD 24-JAN-2002.
 XX
 PF 17-JUL-2001; 2001MO-IL000653.
 XX
 PR 18-JUL-2000; 2000IL-00137345.
 XX
 PR 15-DEC-2000; 2000IL-00140354.
 XX
 PA (COMP-) COMPUGEN LTD.
 XX
 PI Mintz L, Freilich S, Bernstein U;
 XX
 DR WPI; 2002-155037/20.
 XX
 DR N-PSDB; ABB39747.
 XX
 PT One hundred and twenty eight novel nucleic acid sequences, useful for
 PT treating and diagnosing e.g. cancer, asthma and Alzheimer's.
 XX
 PS Claim 6; Page 213-214; 290p; English.
 XX

CC ABB39691 to ABB39818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
 CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
 CC antitubercular, antiarthritic, antipsoriatic, ophthalmological, virucide,
 CC vasoregic, antiarteriosclerotic, antiinflammatory, dermatological,
 CC anorectic, muscular, antinfertilit, hypotension, cardiovascular, cardiant,
 CC anticoagulant, antifibrinolytic, hypotension, cardiovascular, cardiant,
 CC immunomodulator, anticonvulsant, antidiabetic, tranquiliser, antitumor,

CC antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
 CC neurotropic and contraceptive activities. The NS can be used in vaccines,
 CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contraceptive
 XX
 SO Sequence 278 AA;

Query Match 58.0%; Score 1098; DB 5; Length 278;
 Best Local Similarity 99.1%; Pred. No. 1.6e-98;
 Matches 209; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSUSRRKRLDNDLTERPVQKARSGPPRLPCLLPSPPTAPRRAT 60
 Db 29 MRATPLAAPAGSUSRRKRLDNDLTERPVQKARSGPPRLPCLLPSPPTAPRRAT 88
 QY 61 AVATASRLGPPVLLPEEGGRAYQALHCPGTCTYCKVYVQJALAVLEPYARLPKRV 120
 Db 89 AVATASRLGPPVLLPEEGGRAYQALHCPGTCTYCKVYVQJALAVLEPYARLPKRV 148
 QY 121 ARTEVLATGQLIYAFPTTHGMSHIVTRHRIPBEAAVLFRQMTALAHCHQHLV 180
 Db 149 ARTEVLATGQLIYAFPTTHGMSHIVTRHRIPBEAAVLFRQMTALAHCHQHLV 208
 QY 181 RDLKLCRFVADRRKRLVLENDSCVLTG 211
 Db 209 RDLKLCRFVADRRKRLVLENDSCVLTG 239

RESULT 12
 ABB06108
 ID ABB06108 standard; protein; 278 AA.

AC ABB06108;
 DT 10-MAY-2002 (first entry)
 DE Human NS protein sequence SEQ ID NO:200.
 XX
 KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;
 KW antitubercular; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
 KW vasoregic; antiarteriosclerotic; antiinflammatory; dermatological;
 KW anorectic; muscular; antinfertilit; cardiovascular; anticonvulsant;
 KW anticonvulsant; antidiabetic; tranquiliser; immunomodulator; cardiant;
 KW gastrointestinal; virucide; antitumor; antidepressant; neuroleptic;
 KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
 KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
 KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KW infertility; cardiovascular disease; coagulation disease; hypertension;
 KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
 KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KW gastric ulcer; Alzheimer's disease.

OS Homo sapiens.
 XX
 PN WO200206315-A2.
 XX
 PD 24-JAN-2002.
 XX
 PF 17-JUL-2001; 2001MO-IL000653.
 XX
 PR 18-JUL-2000; 2000IL-00137345.
 XX
 PR 15-DEC-2000; 2000IL-00140354.
 XX

PA (COMP-) COMPUGEN LTD.
XX
XX Minter L, Freilich S, Bernstein J;
XX
XX MPI: 2002-155037/20.
DR N-PSDB; ABL39762.
PT One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's.
XX
XX
PS Claim 6; Page 231-232; 290pp; English.
XX
XX ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virologic,
CC vasotrophic, antiarteriosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antifertility, cardiovascular, cardiac,
CC anticonvulsant, antifibrinolytic, hypotension, antidiabetic, tranquiliser, antitumor,
CC immunomodulator, anticonvulsant, antiepileptic, cerebroprotective,
CC neuroprotectant, gastrointestinal, neuroleptic, cerebroprotective,
CC nociceptive and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive
XX
SQ Sequence 278 AA;
Query Match 58.0%; Score 1098; DB 5; Length 278;
Best Local Similarity 99.4%; Pred. No. 1,6e-98;
Matches 209; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MRATPLAAPAGSLSRKKRLLELDNLTDPVQKARSGPOPLPCLLPSPPTAPDRAT 60
DB 29 MRATPLAAPAGSLSRKKRLLELDNLTDPVQKARSGPOPLPCLLPSPPTAPDRAT 88
QY 61 AVATASRLGPVYLLEPEBEGRAYQALHCPGTETGKVPVQALAVLEPYARLPKHAY 120
DB 89 AVATASRLGPVYLLEPEBEGRAYQALHCPGTETGKVPVQALAVLEPYARLPKHAY 148
QY 121 ARPTVLVAGTQLIYAFTTRTHGMSLVTRHRIPEPEAAVLFROMATALAHCHQGLV 180
DB 149 ARPTVLVAGTQLIYAFTTRTHGMSLVSPHRIPEPEAAVLFROMATALAHCHQGLV 208
QY 181 RDLKLCRFVAFDRERKKLVLENLEDSCVLTG 211
DB 209 RDLKLCRFVAFDRERKKLVLENLEDSCVLTG 239
RESULT 13
ADD14087
ID ADD14087 standard; protein; 343 AA.
XX
XX ADD14087;
XX
XX 01-JAN-2004 (first entry)
XX
XX Human src biomarker polypeptide SEQ ID NO:276.
XX
XX predictor set; protein tyrosine kinase activity modulator;
KM protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KM gene therapy; drug sensitivity; genetic profile; cancer; human.
XX
OS Homo sapiens.
XX
XX W02003062395-A2.

XX
XX 31-JUL-2003.
XX
XX 17-JAN-2003; 2003WO-US001981.
XX
XX 18-JAN-2002; 2002US-0350061F.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Huang F, Fairchild CR, Lee FY, Shaw P,
XX
XX MPI: 2003-636735/60.
XX
XX N-PSDB; ADD14683.
XX
XX New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.
XX
XX
PS Claim 10; SEQ ID NO 276; 139pp; English.
XX
XX The present invention describes a predictor set comprising a plurality of
CC polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 343 AA;
Query Match 39.8%; Score 753.5; DB 7; Length 343;
Best Local Similarity 47.3%; Pred. No. 9,9e-65;
Matches 167; Conservative 50; Mismatches 115; Indels 21; Gaps 6;
QY 2 RAPTPLAAPAGSLSRKKRLLELDNLTDPVQKARSGPOPLPCLLPSPPTAPDRAT 61
DB 5 RAPTPLAAPAGSLSRKKRLLELDNLTDPVQKARSGPOPLPCLLPSPPTAPDRAT 57
QY 62 VATASRLGPVYLLEPEBEGRAYQALHCPGTETGKVPVY--QEALAVLEPYARLPKH 118
DB 58 V--SCIGKYLLEPLEGSHVFRVAHLHSGEELVCKVPDISCYQESIA--PCFCIAHS 111
QY 119 HVARPTVLVAGTQLIYAFTTRTHGMSLVTRHRIPEPEAAVLFROMATALAHCHQGL 178
DB 112 NINQITEIILIGETKAVVFFERSYGDHNSFVETCKKUREEPAALFFQIASAVAHCHDGL 171
QY 179 VARDKLCRFVAFDRERKKLVLENLEDSCVLTGPDLSIMPKHACPAVVGCEIISRSAYS 238
DB 172 VARDKLCRFVAFDRERKKLVLENLEDSCVLTGPDLSIMPKHACPAVVGCEIISRSAYS 231
QY 239 GKADVWSLGLVALFTMLAGHYPPQDSEPVLLFGKIRRGAYALPAGLSAPARCTVRCLLR 298
DB 232 GKADVWSLGLVALFTMLAGHYPPQDSEPVLLFGKIRRGAYALPAGLSAPARCTVRCLLR 291

QY 299 EPAERLTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVPDGLGDEAREERGD 351
 DB 292 EFSRLTSQETLHPMPSTID-----FVSNSNAGAKVESDQVDPVMEENLD 339

RESULT 14
 ABB80975
 ID ABB80975 standard; protein; 372 AA.
 AC ABB80975;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human tribbles homologue-1 (htrb-1) polypeptide.
 XX
 KW Human; tribbles; htrb-1; stress kinase inhibitor protein; SKIP-1; Ap-1;
 KW antithrombotic; antithrombotic; antidiabetic; antiproliferative; osteopathic;
 KW ophthalmological; cardiac; cytosolic; haemostatic; immunosuppressive;
 KW antiinflammatory; estrogen receptor; fibroblast growth factor; FGF;
 KW tumour necrosis factor; TNF.
 XX
 OS Homo sapiens.
 XX
 PN MO200253743-A2.
 XX
 PD 11-JUL-2002.
 XX
 PF 08-JAN-2002; 2002WO-US000070.
 XX
 PR 08-JAN-2001; 2001US-0260294P.
 XX
 PA (INTE-) INTERLEUKIN GENETICS INC.
 XX
 PI Dower S, Qianstrom E, Kiss-Toth E;
 XX WPI; 2002-590635/63.
 DR N-PSDB; ABB86478.
 DR
 XX
 PT Novel isolated human tribbles homologue-1 polypeptide for inhibiting Ap-1
 PT mediated inflammatory signal in a cell, and activating Akt-mediated
 PT signal e.g. Ap-1-mediated gene activation signal in a cell.
 XX
 PS Claim 21; Fig 10B; 131pp; English.
 XX
 CC The invention provides an isolated human tribbles homologue-1 (htrb-1,
 CC also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
 CC htrb polypeptide is useful for inhibiting an Ap-1 mediated inflammatory
 CC signal in a cell. The polypeptide employed in the method is preferably
 CC htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N htrb-3 C, or htrb-
 CC 3 N C. It is also useful for providing htrb agonist activity for
 CC activating an ERK-mediated signal e.g. Ap-1-mediated gene activation
 CC signal, an estrogen receptor-mediated gene activation signal, an
 CC fibroblast growth factor (FGF) induced signal, or a PMA induced signal,
 CC in a cell. Htrb modulators are useful for modulating Ap-1 mediated
 CC inflammatory signal in a cell such as tumor necrosis factor (TNF) induced
 CC inflammatory signal, or an interleukin induced inflammatory signal. htrb
 CC proteins are useful in screening assays, predictive medicine and in
 CC therapeutics or prophylactics. The htrb proteins are useful for screening
 CC compounds e.g. for treating and/or preventing diseases caused by abnormal
 CC htrb activity, such as rheumatoid arthritis, diabetes, psoriasis,
 CC osteoporosis, diabetic retinopathy, myocardial infarction and cancers.
 CC The htrb therapeutics are useful for antagonizing interleukin-1 dependent
 CC disorders of human placenta, intraventricular hemorrhage, neonatal white
 CC matter damage and subsequent cerebral palsy, and inflammation or
 CC autoimmune disorders. The present sequence represents the htrb-1
 CC polypeptide
 XX
 SQ Sequence 372 AA;

Query Match 39.0%; Score 737.5; DB 5; Length 372;
 Best Local Similarity 46.2%; Pred. No. 4,1e-63;
 Matches 160; Conservative 49; Mismatches 126; Indels 11; Gaps 6;

QY 2 RATPLAAPA-GSLSRKRLLEID--NIDTERPVQKRRSGP--QPIPLCLPLSPPTA 55
 DB 17 RGPALLPFARRGVPAKRLPADDAVAACCPRLSECCSSPPDYUSPPGSPC-SPPGPPAA 75

QY 56 PDRATAVATA---SRLEPVYLLEPEGGRAYVQALHCPGTETGYTKVYPVQALAVLEPPYA 112
 DB 76 PGAGGGSGSAPGFSRLADYLLPLPAERHVSRLCHTGRRLKRVKVPPIKHYQDKIRPYI 135

QY 113 RLPPKHVAPPTVEYLAQTOLLYAEFTTRHGDMSLVTRHRIPEPEAAVLFRQWATALAH 172
 DB 136 QLPFSHNSITGIVEYIIAGETKAYVFEKSPGDMHSYVARSRLREKREAAARLFQIVSAVAH 195

QY 173 CHQHGVLRLPLCKRFPAPRERKVLVLENDSCVITGPPDSCIMDYACPAYGPIILS 232
 DB 196 CHQSAIVLGDCLKRKFFVSTEEFTQLTESLEDTHIMKGBDADLSDHGCPAYVSPRIIN 255

QY 233 SRASVSGKADVMSLGVALLFTMLAGHYPPFODSEPVLLFGKIRGVALPAGLSAPARCIV 292
 DB 256 TTGTYSKADVMSLGVMLTYTLVGRYPFHDSDPSALFSKIRKQPCIPHHISPKARCLI 315

QY 293 RCLRRRPAERLTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVPD 338
 DB 316 RSLRRRPSERLTAPETLLHPMFESVLEP-GYIDSEIGTSDQIYPE 360

RESULT 15
 AAU28135
 ID AAU28135 standard; protein; 269 AA.
 XX
 AC AAU28135;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secretory protein, Seq ID No 304.
 XX
 KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.
 XX
 OS Homo sapiens.
 XX
 PN MO200166689-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 05-MAR-2001; 2001WO-US004942.
 XX
 PR 07-MAR-2000; 2000US-00519705.
 PR 19-MAY-2000; 2000US-00574454.
 PR 17-JUN-2000; 2000US-00596193.
 PR 14-JUL-2000; 2000US-00616847.
 PR 19-SEP-2000; 2000US-00665363.
 PR 20-OCT-2000; 2000US-00693267.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 DR N-PSDB; AAS45035.
 DR
 XX
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders.
 XX

PS Example 4; SEQ ID NO 304; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I) and polynucleotides (II). (I) and (II) are useful for treating inflammatory conditions such as arthritis, nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis, immune responses, and is involved in increasing haematopoiesis, stem cell survival, bone growth and remodeling. (I), (II) and modulators of (II) are useful for prophylaxis or treatment of one or more cancers. (II) is also useful for creating transgenic animals useful for studying the in vivo activities of the polypeptide as well as for studying modulators of the polypeptides. (I) induces the proliferation of neural cells and regeneration of nerve and brain tissue and is useful for the treatment of central and peripheral nervous system diseases and neuropathies, such as Alzheimer's, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis. In addition, (II) is involved in chemotactic or chemokinetic activity, regulation of haematopoiesis and is useful for treating myeloid or lymphoid cell disorders, platelet disorders such as thrombocytopenia and for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue growth, and in tissue repair, healing of burns, incisions, ulcers, for treating osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease. Furthermore, (I) is also useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, various immune deficiencies and disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic reactions and conditions, such as asthma or other respiratory problems. In addition, (I) affects biorhythms or circadian cycles of rhythms, fertility, metabolism, catabolism, anabolism, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides analgesic effects or other pain reducing effects, immunoglobulin like activity and can act as an antigen in a vaccine composition to raise an immune response. AAU28020-AAU28395 represent novel human secreted protein amino acid sequences of the invention

CC
XX
SQ Sequence 269 AA;

Query Match 35.9%; Score 678.5; DB 4; Length 269;

Best Local Similarity 51.4%; Pred. No. 1.5e-57; Mismatches 132; Conservative 39; Matches 79; Indels 7; Gaps 2;

QY 88 CP-----TGETYCKVYPVOEALAVLEPYARLPKHVARPTEVLAGTQLVAFPTRTH 141
DB 2 CPERCASTLGRRCVCFPIKHVQDKIRPYTLQPSHNTIGIVEVILGETKAYVFEKDF 61
QY 142 GDMHSLVTRTHRIPEPPAAVLPRQMATALAHQHGIVLRDLKLCRFVFPADRRKCLVE 201
DB 62 GDMHSYVRSRKRLEEEAARLFQOIVSAVACHQSAIVLGDILKRFKVFSTEERTQLRLR 121
QY 202 NLEDSCVLGGPDSIMWCKHCPAYVGPETLSRASYSAGADVMSLGVALFTMLAGHYPF 261
DB 122 SLBDTHMKGEDALSDKSGCPANVSPELINTGTYSCKADVMSLGVMLYTLVGRYPF 181
QY 262 QDSEPVLLFGKIRGAYALPAGLSAPARCLVRCILRREPARELTATGILLHPWLRODMP 321
DB 182 HDSDPSALFVKIRGQCFCEHHSPPKARCILRSILRREPSERLTAPETILLHPWESVLEP 241
QY 322 LAPTRSHLMEAAQVVPD 338
DB 242 -GYIDSEIGTSQIVPE 257

Search completed: August 24, 2004, 18:52:13
Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 18:51:12 ; Search time 18 Seconds
(without alignments)
1026.783 Million cell updates/sec

Title: US-10-070-337-5
Perfect score: 1892
Sequence: 1 MRATPLAAPAGSLSRKKRL.....GLGLDEARBEEGDRVLYG 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep: *
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep: *
3: /cgn2_6/prodata/2/iaa/5A_COMB.pep: *
4: /cgn2_6/prodata/2/iaa/5B_COMB.pep: *
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep: *
6: /cgn2_6/prodata/2/iaa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1888	99.8	358	4	US-09-799-875-8 Sequence 8, Appl
2	1676	88.6	360	4	US-09-509-802A-9 Sequence 9, Appl
3	291.5	15.4	607	2	US-08-878-989-15 Sequence 15, Appl
4	291.5	15.4	607	3	US-09-272-796-15 Sequence 15, Appl
5	273.5	14.5	633	3	US-08-557-006C-43 Sequence 43, Appl
6	272.5	14.4	668	4	US-09-930-181-2 Sequence 2, Appl
7	267.5	14.1	252	2	US-07-857-224B-26 Sequence 26, Appl
8	264.5	14.0	339	3	US-08-688-988-33 Sequence 33, Appl
9	262	13.8	257	2	US-07-857-224B-25 Sequence 25, Appl
10	261.5	13.8	264	2	US-07-857-224B-19 Sequence 19, Appl
11	259.5	13.7	552	3	US-08-557-006C-40 Sequence 40, Appl
12	257	13.6	504	4	US-09-554-726A-10 Sequence 10, Appl
13	253	13.4	354	3	US-08-688-988-29 Sequence 29, Appl
14	249.5	13.2	433	2	US-08-913-050A-7 Sequence 7, Appl
15	249.5	13.2	433	2	US-08-749-902-5 Sequence 5, Appl
16	249	13.1	264	2	US-07-857-224B-18 Sequence 18, Appl
17	248.5	13.1	685	2	US-08-878-989-1 Sequence 1, Appl
18	248.5	13.1	685	3	US-09-136-282-2 Sequence 2, Appl
19	248.5	13.1	685	3	US-09-272-796-1 Sequence 1, Appl
20	248.5	13.1	685	3	US-09-505-744-2 Sequence 2, Appl
21	247.5	13.0	382	3	US-09-142-551A-2 Sequence 2, Appl
22	246.5	13.0	353	3	US-08-688-988-31 Sequence 31, Appl
23	246.5	13.0	504	4	US-09-554-726A-14 Sequence 14, Appl
24	245	12.9	359	3	US-08-688-988-32 Sequence 32, Appl
25	243.5	12.9	358	4	US-09-230-896C-29 Sequence 29, Appl
26	243.5	12.9	556	4	US-09-800-960-4 Sequence 4, Appl
27	243.5	12.9	556	4	US-10-096-960-4 Sequence 4, Appl

ALIGNMENTS

28	243.5	12.9	565	4	US-09-800-960-2	Sequence 2, Appl
29	243.5	12.9	565	4	US-10-096-960-2	Sequence 2, Appl
30	243.5	12.9	604	4	US-09-523-849-35	Sequence 35, Appl
31	242.5	12.8	603	4	US-09-930-181-17	Sequence 17, Appl
32	241.5	12.8	351	3	US-08-688-988-28	Sequence 28, Appl
33	241	12.7	295	1	US-07-951-715A-23	Sequence 23, Appl
34	241	12.7	295	2	US-08-459-448A-23	Sequence 23, Appl
35	241	12.7	295	3	US-08-459-595A-23	Sequence 23, Appl
36	241	12.7	295	3	US-08-459-504B-23	Sequence 23, Appl
37	241	12.7	295	3	US-08-459-444-23	Sequence 23, Appl
38	241	12.7	295	4	US-09-547-422-23	Sequence 23, Appl
39	241	12.7	455	4	US-09-554-726A-7	Sequence 7, Appl
40	241	12.7	484	4	US-09-554-726A-5	Sequence 5, Appl
41	239	12.6	345	3	US-09-101-146-1	Sequence 1, Appl
42	238	12.6	273	1	US-08-252-995D-10	Sequence 10, Appl
43	238	12.6	273	1	US-08-834-108-10	Sequence 10, Appl
44	238	12.6	416	1	US-08-252-995D-2	Sequence 2, Appl
45	238	12.6	416	2	US-08-834-108-2	Sequence 2, Appl

RESULT 1
US-09-799-875-8
; Sequence 8, Application US/09799875
; Patent No. 6638721
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: No. 6638721el Human Protein Kinases and Uses
; TITLE OF INVENTION: Therefore
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/09/799, 875
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182, 059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659, 287
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-875-8

Query Match 99.8%; Score 1888; DB 4; Length 358;
Best Local Similarity 99.7%; Pred. No. 8e-186; 0; Indels 0; Gaps 0;
Matches 357; Conservative 1; Mismatches 0

QY	1	MRATPLAAPAGSLSRKKRLDNDLTERPVQKRAASGPQRLPLSPPTAPDRAT	60
DB	1	MRATPLAAPAGSLSRKKRLDNDLTERPVQKRAASGPQRLPLSPPTAPDRAT	60
QY	61	AVATASRLGPVYLLEPEEGRAYQALHCPGTETTCVYVQALAVLEPYARLPKHV	120
DB	61	AVATASRLGPVYLLEPEEGRAYQALHCPGTETTCVYVQALAVLEPYARLPKHV	120
QY	121	ARTEVLVAGQQLYAFPTRTGHMHSIVRRHRIPEEAVALFRQKATLAHCHQGLVL	180
DB	121	ARTEVLVAGQQLYAFPTRTGHMHSIVRRHRIPEEAVALFRQKATLAHCHQGLVL	180
QY	181	RDJLCLCFVADBERKKLVLENLEDSCVLTGPPDSLMDKACPAVVGPEILSSRASYSGK	240
DB	181	RDJLCLCFVADBERKKLVLENLEDSCVLTGPPDSLMDKACPAVVGPEILSSRASYSGK	240
QY	241	AAVWSLGVALLFTMLAGHYPPQSEPYLLFGKTRRGAYALPAGISAPARCLVRLLRREP	300
DB	241	AAVWSLGVALLFTMLAGHYPPQSEPYLLFGKTRRGAYALPAGISAPARCLVRLLRREP	300
QY	301	AERLTATGILLHPRLRQDPMLAPTRSHLWEAQQVVDGLGLDEARBEEGDRVLYG	358

Db 301 AERLTATGILHFWLRODPMPLAPTRSHLMEAAQVVPDGLDERREBEGREVLYG 358

RESULT 2
US-09-509-902A-9
Sequence 9, Application US/09509902A
Patent No. 6387676
GENERAL INFORMATION:
APPLICANT: Virca, Duke
APPLICANT: Bird, Timothy A.
APPLICANT: Anderson, Dirk M.
APPLICANT: Marken, John S.
TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
FILE REFERENCE: 2877-US
CURRENT APPLICATION NUMBER: US/09/509, 902A
CURRENT FILING DATE: 1999-08-03
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 360
TYPE: PRT
ORGANISM: Homo sapiens
US-09-509-902A-9

Query Match 88.6%; Score 1676; DB 4; Length 360;
Best Local Similarity 99.7%; Pred. No. 5, 2e-164;

Matches 317; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSRKRLDLNDLTERPVOKRASGPPRLPCLPLSPPTAPDRAT 60
DB 43 MRATPLAAPAGSLSRKRLDLNDLTERPVOKRASGPPRLPCLPLSPPTAPDRAT 102
QY 61 AVATASRLGYYVLEPEEGRAYOALHCPCTGYTCVYVOALATLEAYALPPKHV 120
DB 103 AVATASRLGYYVLEPEEGRAYOALHCPCTGYTCVYVOALATLEAYALPPKHV 162
QY 121 ARPEVLATGQLLYAFRTTHGDMHSIVTRHRIPEPEAAVLPRQATLAHCHQGLV 180
DB 163 ARPEVLATGQLLYAFRTTHGDMHSIVTRHRIPEPEAAVLPRQATLAHCHQGLV 222
QY 181 RDLKTRFVADDERKKVLVLENDSCVLTGPDSDLMKACPAVYGPETLSSRASYSCK 240
DB 223 RDLKTRFVADDERKKVLVLENDSCVLTGPDSDLMKACPAVYGPETLSSRASYSCK 282
QY 241 AADVMSLGVALLFTMLAGHYFQDSEPVLLPGKTRRGAYALPALSAPARCLVCLRRP 300
DB 283 AADVMSLGVALLFTMLAGHYFQDSEPVLLPGKTRRGAYALPALSAPARCLVCLRRP 342
QY 301 AERLTATGILHFWLROD 318
DB 343 AERLTATGILHFWLROD 360

RESULT 3
US-08-878-989-15
Sequence 15, Application US/08878989
Patent No. 5865803
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl G.
APPLICANT: Lal, Preeti
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto

STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,989
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0321 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 607 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1827450
US-08-878-989-15

Query Match 15.4%; Score 291.5; DB 2; Length 607;
Best Local Similarity 29.2%; Pred. No. 3e-21;
Matches 92; Conservative 45; Mismatches 129; Indels 49; Gaps 12;

QY 48 LPLSP-----TAPDRATAVATASRLGYYVLEPEEGG--RAYOALHCPCTGYTCVYVP 101
DB 5 LPTSDPGRLLTPDRSGRTYLYKGRLLG-----KGFARCYEATDTETGSAVAVKVIFQ 56
QY 102 QEALAVLEBYARLP-----PKHVARPEVLATGQLLYAFTR--THGDMHSIVR 149
DB 57 SR---VAKPHQEKILNELIELRDLQHRHIVRSHPEDADNIIYILELCSKSLAHIMK 113
QY 150 THRRIPEPEAAVLPRQATLAHCHQGLVLRDLKTRFVADDERKKVLVLENDSCVLT 209
DB 114 ARHTLLEPEVRYLYKQILSGLYKHQGLIHRDLKGNFTT--ENMELKVGDFGLAARL 171
QY 210 TGPDSLMDXKAC--PAYVGPETLSSRASYSCKAADVMSLGVALLFTMLAGHYFQDSEPV 267
DB 172 EPEEGR--KRTICGTPNYVAPVLLRQG--HGFEADVMSLGVCMYTLTGSPPEFETADLK 227
QY 268 LTFGKIRGAYALPGLSAPARCLVCLRRPARELTATGILHFWLROD---DMPILA 323
DB 228 EYRRCIKQVHYTLPSLSLPARQILAALIRASPRDPSIDQILRHDFFTKGYTPDRPLTS 287
QY 324 PTRSHLMEAAQVVPD 338
DB 288 -----SCVTVPD 294

RESULT 4
US-09-272-796-15
Sequence 15, Application US/09272796
Patent No. 6207148
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl G.
APPLICANT: Lal, Preeti


```

; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purni
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,796
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/878,989
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 607 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1827450
; US-09-272-796-15

```

Query Match 15.4%; Score 291.5; DB 3; Length 607;
 Best Local Similarity 29.2%; Pred. No. 3e-21;
 Matches 92; Conservative 45; Mismatches 129; Indels 49; Gaps 12;

```

QY 48 LPLSP---TAPDRATVATASRLGPGVYLLEPEBGG--RAYQALHCPGTGTEYTCVYV 101
DB 5 LFTSDGRLITDPRSGRTYLGRLG-----KGGFARCYEATDTETGSAVAVV 56
QY 102 QELANLEPYARL-----PHKIVARPEVLAGTQLYAFPTT--THGDMSLVR 149
DB 57 SR---VAKPHQRKILNELHRLDQHRHVRFSHHFEDADNIYILELCSRSLSAHYK 113
QY 150 TRHRIPEPEAAVLFRQMATALAHQHGLVLRDLKCRFEVADREKRLVLENLDSCVL 209
DB 114 AKTITLPEPEVRYVIRKQILSGILKRGILHRLDKGNFTT--EMNELKVGDFGAARL 171
QY 210 TGPDDSLMDKAC--PAYVGPETLSSRASYSKADVMSIGVALFTMLAGHYFPQSEPV 267
DB 172 EPPEQR--KKTICGTENVYVAVLIRQG--HGEADVWSIGCVMYTLLCGSPFETADLK 227
QY 268 LLEFGKIRGAYALPAGISAPARCLVRCILRREPAERLTATGILHPMLRQ---DMPPLA 323
DB 228 EYTRCIRKQVYVTPASISLPARQLALIRASPRDRPSIDQILRHDFFTGYTPDLPLIS 287
QY 324 PTRSHLMEAAQVVPD 338
DB 288 -----SCVTVPD 294

```

RESULT 5

```

US-08-557-006C-43
; Sequence 43, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
; APPLICANT: Bert, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forder, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NGAP/PMH37588/UST
; CURRENT APPLICATION NUMBER: US/08/557,006C
; PRIOR FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 43
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Yeast
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(633)
; OTHER INFORMATION: Yeast SNF1 polypeptide
US-08-557-006C-43

```

Query Match 14.5%; Score 273.5; DB 3; Length 633;
 Best Local Similarity 27.8%; Pred. No. 2.2e-19;
 Matches 88; Conservative 53; Mismatches 138; Indels 37; Gaps 11;

```

QY 51 SEPTAPDRATVATASRLGPGVYLLEP-BEG--GRAYQALHCPGTGTEYTCVYVQELAV 107
DB 35 SNSTLNKSSLDAGAHGNVQIVKTLGBSGKVKLAHHTTGQKVALKIN--KKVLAK 93
QY 108 IE-----PYALLPHKIVARPEVLAGTQLYAFPTT--THGDMHSLVRHRIPEPE 158
DB 94 SDMGRIERELSYLRLHPIHLIKLDYVTKSDEILMIVYAGNELFDYIVQDRKSEOE 153
QY 159 AAVLFRQMATALAHQHGLVLRDLKCRFEVADREKRLVLENLDSCVLTCGDDSLMD 218
DB 154 ARPFQQTISAVEYCHRKIVYHRLKPEMLLDEHINVKIAPGL--SNIMT---DGNFL 208
QY 219 KHAC--PAYVGPETLSSRASYSKADVMSIGVALFTMLAGHYFPQSEPVILFGKIRRG 276
DB 209 KNSCGSPNVAARPEVLSGKL--YAGPEVDVMSCGVILYVMLCRRLPFDDISIPVLFKNISNG 267
QY 277 AYALPAGISAPARCLVRCILRREPAERLTATGILHPMLRQ--PMFLAPTRSHLMEAAQV 335
DB 268 VYTLKRFSLPGAAGLIKRLIVNPLNRSIHIMODDWRKVDLPEYLP----- 316
QY 336 VPDGLGLDEAREEGD 351
DB 317 -PD---LKPHEBENE 328

```

RESULT 6
 US-09-930-181-2
 ; Sequence 2, Application US/09930181
 ; Patent No. 6455292
 ; GENERAL INFORMATION:
 ; APPLICANT: Origene Technologies
 ; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
 ; FILE REFERENCE: 16U 101 V1
 ; CURRENT APPLICATION NUMBER: US/09/930,181
 ; CURRENT FILING DATE: 2001-08-16
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: Patentln version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 668
 ; TYPE: PRT

[illegible]

```

RESULT 7
US-07-857-224B-26
; Sequence 26, Application US/07857224B
; Patent No. 5958784
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/857,224B
; FILING DATE: 03/25/92
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA: none
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (International) 41 1 632 2830
; TELEFAX: (International) 41 1 262 2437
; TELEX: none
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION:
; ORIGINAL SOURCE:
; ORGANISM: Schizosaccharomyces pombe
; FEATURE: Protein kinase; Table 8 Column 29
; PUBLICATION INFORMATION:
; AUTHORS:

```

? AUTHORS: Hankins, S. K.
? AUTHORS: Quinn, A. M.
? AUTHORS: Hunter, T.
? TITLE: The protein kinase family
? JOURNAL: Science
? VOLUME: 241
? PAGES: 43-52
? DATE: 1968
US-07-857-224B-26

Query Match	14.1%;	Score 267.5;	DB 2;	Length 252;
Best Local Similarity	31.5%;	Pred. No. 2.4e-19;		
Matches 74;	Conservative 44;	Mismatches 108;	Indels 9;	Gaps 6

QY 85 ALHOPCTGEYCYCKYYPVOEALVALE--PYALRPHKHVAPRPVYLAGTOLLV--AEPTRTHG 142

Db 20 AKAAKTGDLAAIKIIPIRYASIGHEILMMRLIRHNILRLVDWTDHDMHMLAEYVBDG 79

QY 143 DMHSIVTRHRIRPPEAAVLFROMATALAHCHOGLVLRDYLCPFYADREKRKLVDEN 202

Db 80 ELFYHIRKHGPLSRERAAHYLSQLLDVAVHCHRRFRFRIRDKLENILIKVNEQC--IKI 136

QY 203 LEBSCVLTGPDDSLMDYACPA--YVGDEILSSPASYSGKADWYSLGVALFTMLAGHYR 260

Db 137 ADFMGATVEPDDSCLENY--CGSLHYLPAEIVSHK--PYRGADADWYSCGVILYSLSNLCP 194

QY 261 FQDESEVLFGKIPRGAYVALPAGISAPARCIVRCLLREPRERLTATGILHFMVL 315

Db 195 FGGCNTVIVYINKIRHGVADLPSSSTSSAODLHMLLDVNSTRTLTIEEFESHPTL 249

```

RESULT 8
US-08-688-988-33
; Sequence 33, Application US/08688988B
; Patent No. 6095545
;
; GENERAL INFORMATION:
; APPLICANT: Lefebvre, Daniel D.
; APPLICANT: Malboobi, Mohammad A.
; TITLE OF INVENTION: PHOSPHATE STARVATION-INDUCIBLE PROTEINS
; FILE REFERENCE: PPL96-03
; CURRENT APPLICATION NUMBER: US/08/688,988B
; CURRENT FILING DATE: 1996-07-31
; NUMBER OF SEQ. ID NOS.: 48
; SOFTWARE: PasteSeq for Windows Version 3.0
; SEQ. ID NO. 33
;
; LENGTH: 339
;
; TYPE: PRT
;
; ORGANISM: Glycine max
; US-08-688-988-33

```

	Query Match	14.0%: Score 264.5; DB 3; Length 339;
	Best Local Similarity	28.7%: Pred. No. 7.6e-19;
	Matches	86; Conservative 34; Mismatches 87; Indels 93; Gaps 12;
QY	117 HKHVAAPLEVLATGTLIYAFPTRT	-----GDMSLVTRRRIPEPEAAVLFROM 166
Db	58 HNNIRKEV-----FLTPHTALVLEAAGGLFERRI	CNAGLSDEAFAFFQOL 108
QY	167 ATAAHCHQGLVRLDKL-----CRVPADRREKKVL	ENLDESCVLTG 211
Db	109 ISGVSYCHSQMIGCRDILKENTLIDGNPAIRLIK	CDPGRS-----KSLMLISQPKSYVGT- 163
QY	212 PDSLMWKACPAVYGEIILSRASYS	GKADWLSGLVALFTMLAGHYPFQDSBPVILF- 270
Db	164 -----PAAIAPEVL-SREKEYGKADVWSC	CVTLTYVWLVCAYPFPEDEEDPKNFR 211
QY	271 ---GKIRGAYALP--AGLSAPARCIVRC	LIREPARELTATGILLHPWTRQD--PMDL-- 322
Db	212 KSIQRINSVQYALPDYRVYSKECRHLIS	RFVNPAPRINISIKQLMWRKMLPREITE 271
QY	323 -----APTRSLMEAAQVVPDGLGL-DEAREE	BGDEY 354
Db	272 AERRGVEETKQDPQSGVEEIMQII	OEARKIKHTGSA-----GTGTSIVNRCDENAEV 326

Db 203 FWEDQKLYQOIKAGAYDPSPEDMTVTEAKNLIQMLTINPAKTIHAEALKEHWPVC 262
 QY 317 Q 317
 Db 263 Q 263

RESULT 11
 US-08-557-006C-40
 ; Sequence 40, Application US/08557006C
 ; Patent No. 6258547
 ; GENERAL INFORMATION:
 ; APPLICANT: Berl, Rajindar K.
 ; APPLICANT: Carling, David
 ; APPLICANT: Forde, Robert A.
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
 ; FILE REFERENCE: NGAP/PHM37588/UST
 ; CURRENT APPLICATION NUMBER: US/08/557,006C
 ; CURRENT FILING DATE: 1996-03-06
 ; PRIOR APPLICATION NUMBER: PCT/GB94/01093
 ; PRIOR FILING DATE: 1994-05-20
 ; PRIOR APPLICATION NUMBER: GB 9310489.1
 ; PRIOR FILING DATE: 1993-05-21
 ; PRIOR APPLICATION NUMBER: GB 9318010.7
 ; PRIOR FILING DATE: 1993-08-31
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 40
 ; LENGTH: 552
 ; TYPE: PRT
 ; ORGANISM: Rat
 ; FEATURE:
 ; NAME/KEY: gene
 ; LOCATION: (1)-(1747)
 ; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
 ; OTHER INFORMATION: Fragment Begins at nucleotide 24 and ends with
 ; OTHER INFORMATION: nucleotide 1765
 ; US-08-557-006C-40

Query Match 13.7%; Score 259.5; DB 3; Length 552;
 Best Local Similarity 29.2%; Pred. No. 5e-18;
 Matches 82; Conservative 43; Mismatches 119; Indels 37; Gaps 9;

QY 67 RLGPYVL--LEPESGRAYQALHCPTEYCKVYVQE--ALAVLBPYAR-----LP 115
 Db 12 KIGHVYLGDTLGIGTGKXKIGEHQLTGHKAVAKIINRQKIRSLDVVGKIKRBIQNLKLF 71
 QY 116 PKHVARPTVEVA-GTQLIYAFPTRTGDMHSLVTRHRIPEEAAVLFRQMATALAHCH 174
 Db 72 RHPHIKLYQVISTPDPFPMVMEVSGELFDYICKHGRVEVEARRLFQQLISADVCH 131
 QY 175 QHGLVLRDLKCRFVADREKRLV---LEN-----LEDSCVLTPDDSLMDKACCA 224
 Db 132 RIMVVRHDLKPENVILDAQMAKIALDFGSLNMSDGEFRTSC-----GSPN 178
 QY 225 YVGPETLSRASYSQKADWVSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAQL 284
 Db 179 YAAPEVISGRLL-YAGPEVDIWSOGVILVALLCGTLPFDEHNPITLFFKIRGGVFIPEYL 237
 QY 265 SAPARCIVRCILRREPAERILTGTLLHPMLRQD-PMPLAP 324
 Db 238 NRSIATLLMHMLQVDPDLKRAIKIKDIRHEHWFQDLPISYLF 278

RESULT 12
 US-09-554-726A-10
 ; Sequence 10, Application US/09554726A
 ; Patent No. 6642369
 ; GENERAL INFORMATION:
 ; APPLICANT: HERRMANN, Bernhard
 ; APPLICANT: KOSCHORZ, Birgit
 ; APPLICANT: KISPERT, Andreas

; TITLE OF INVENTION: NUCLEIC ACIDS INVOLVED IN THE RESPONDER PHENOTYPE AND APPLICATION
 ; FILE REFERENCE: THEREOF
 ; CURRENT APPLICATION NUMBER: US/09/554,726A
 ; CURRENT FILING DATE: 2000-05-18
 ; PRIOR APPLICATION NUMBER: PCT/EP 98/07395
 ; PRIOR FILING DATE: 1998-11-18
 ; PRIOR APPLICATION NUMBER: EP 98 10 3596.7
 ; PRIOR FILING DATE: 1998-03-02
 ; PRIOR APPLICATION NUMBER: EP 97 12 0190.0
 ; PRIOR FILING DATE: 1997-11-18
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 10
 ; LENGTH: 504
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-554-726A-10

Query Match 13.6%; Score 257; DB 4; Length 504;
 Best Local Similarity 28.7%; Pred. No. 8e-18;
 Matches 87; Conservative 48; Mismatches 122; Indels 46; Gaps 13;

QY 71 YVLLPE-EEGGRAYQAL--HCPYGT-----EYTC-KVYPVOEALAVLEPYARLPP 116
 Db 28 YVWLFTIGHGGCATVLAQHRLTGTHVAVKTRKRYWCNRVISEVLLMAD----- 80
 QY 117 HKHVARPTVLAQTQLIYAFPTRTG-DMSLVTRHRIPEEAAVLFRQMATALAHCHQ 175
 Db 81 HPMIISLVQVIEIKKKVYLIMELCKGSKSLYQHIRKAGVLOEHEARALFQQLSAMYCHN 140
 QY 176 HGLVLRDLKCRFVADREKRLVLENEDSCVLTPDDSLMDKACPAY--VGPETLS 233
 Db 141 QGIVHRDLKPDN-IMVEKQKXKIIDFGLTQKVPQKLNLF---CGYFESAPRVLIS 195
 QY 234 RASYSKADWVSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAQLSAPACIVR 293
 Db 196 -IPYDGPRTDWTGLGVLYFMVYTGKIPFDACSIKLVKYLAKGSIPEKLSHEDLLS 254
 QY 294 CLRRPEARLTATGILLHPMLRQ-----DP---MPLAP-----TRSHLMEAAQVVP 337
 Db 255 LMTANPKLRPTVAEVMVHPVWTEGSGVPPDCEEGTPIKPDPAIVKAMGHIGFQADIE 314
 QY 338 DGL 340
 Db 315 DSL 317

RESULT 13
 US-08-688-988-29
 ; Sequence 29, Application US/08688988B
 ; Patent No. 6096545
 ; GENERAL INFORMATION:
 ; APPLICANT: Lefebvre, Daniel D.
 ; APPLICANT: Malboobi, Mohammad A.
 ; TITLE OF INVENTION: PHOSPHATE STARVATION-INDUCIBLE PROTEINS
 ; FILE REFERENCE: PPL96-03
 ; CURRENT APPLICATION NUMBER: US/08/688,988B
 ; CURRENT FILING DATE: 1996-07-31
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 29
 ; LENGTH: 354
 ; TYPE: PRT
 ; ORGANISM: Brassica napus
 ; US-08-688-988-29

Query Match 13.4%; Score 253; DB 3; Length 354;
 Best Local Similarity 29.1%; Pred. No. 1.2e-17;
 Matches 85; Conservative 33; Mismatches 104; Indels 70; Gaps 11;
 QY 117 HKHVARPTE-VIAGTQLIYAFPTRTGDMHSLVTRHRIPEEAAVLFRQMATALAHCHQ 175
 Db 117 HKHVARPTE-VIAGTQLIYAFPTRTGDMHSLVTRHRIPEEAAVLFRQMATALAHCHQ 175

Db 57 HPIIRREKVVLTPTHTLAIMEYAGGELFERICSGRFSSEDEARYFQOLISGVSYCHA 116
QY 176 HGLVLRDLKLT-----CRFVFADREKKLVLENLEDSCVLTGPDSDLMDKH 220
Db 117 MOICRDLKLENTLLDGPAPRLKICDGYG---KSLLSRPSSTVGT-----162
QY 221 ACPAVGPELISRSRYSKADAVSLGVALFTMLAGHYFPDSEPVLLF---GKIRRG 276
Db 163 --PAVIAPEVL--SRREYDGNADVWSCVTLVYMLVGAYPFEDCEDPKGRFKTIQKIMAV 219
QY 277 AYALP--AGLSAPACVLCILRRPARLRTATGLLHPW--LRQDPMPLAPTRSLM---330
Db 220 QYKIPDYVHISQDCCHLSRIEVANSLSRKITTAETKHPWELKNLPRELTETAQAAYKK 279
QY 331 -----EAAQVDP-----GLGH-----DEAREEGDREVV 355
Db 280 ENPTSPQTAETIMKIVDDAKTPPVSRSIGSGFGMGKGDDEEEVEDEEVV 331

RESULT 14
US-08-913-050A-7
Sequence 7, Application US/08913050A
Patent No. 5827726

GENERAL INFORMATION:
APPLICANT: NEZU, Jun-ichi
TITLE OF INVENTION: DNA ENCODING PROTEIN KINASE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 7th Street N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913, 050A
FILING DATE: 05-SEP-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 57104/1995
FILING DATE: 16-MAR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP PCT/JP96/00660
FILING DATE: 15-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: NEZU=4
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-913-050A-7

Query Match 13.2%; Score 249.5; DB 2; Length 433;
Best Local Similarity 27.0%; Pred. No. 3.8e-17;
Matches 69; Conservative 43; Mismatches 129; Indels 15; Gaps 4;

QY 117 HKHVARTEVLAG--TQLLYAFTRTGDMHSLVRT--RRRIPEPAALVFRQMATATAAH 172
Db 107 HKNVITQLVDVLYNEKQKMTWMEYCVCGQEMLDSPVKRFVQCQHGFCQIDLEY 166
QY 173 CHQGLVLRDLKLCRFVFADREKKLVLENLEDSCVLTGPDSDLMDKHACPAYVGPILS 232

Db 167 LHSQIVHKDITKIPENLLITTTGCTIKISDLGVAEALHPAADTORTSGSAPFQPEELAN 226
QY 233 SRASYSKADAVWSLGVALLFTMLAGHYFPDSEPVLLFGKIRGAYALPAGLSAPARCLV 292
Db 227 GLDTFSGFKVDIWSAGVTLVNTTGLVPEFGDNIYKLFENIGKSYALPGDCGPFLLDIL 286
QY 293 RCLIRREPAEKLITATGILLHFWLQD-----PMLAPTRSLHMEAAQVPP--DGLG 341
Db 287 KGMLEYEPARKRSIRQIRQSHMFRKKQPPAPAPVPIPPSPDKORWRSMVTPYLEDLHG 346
QY 342 LDEAREEGDREVVLY 357
Db 347 ADEDEDLFDIEDDITY 362

RESULT 15
US-08-749-902-5
Sequence 5, Application US/087499902
Patent No. 5985635

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN SERINE/THREONINE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,902
FILING DATE: Filed Herewith
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0150 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1480861
US-08-749-902-5

Query Match 13.2%; Score 249.5; DB 2; Length 433;
Best Local Similarity 27.0%; Pred. No. 3.8e-17;
Matches 69; Conservative 43; Mismatches 129; Indels 15; Gaps 4;

QY 117 HKHVARTEVLAG--TQLLYAFTRTGDMHSLVRT--RRRIPEPAALVFRQMATATAAH 172
Db 107 HKNVITQLVDVLYNEKQKMTWMEYCVCGQEMLDSPVKRFVQCQHGFCQIDLEY 166
QY 173 CHQGLVLRDLKLCRFVFADREKKLVLENLEDSCVLTGPDSDLMDKHACPAYVGPILS 232

Mon Aug 30 09:23:51 2004

us-10-070-337-5.rai

Page 8

```

Dh 167 LHSQGI VNHDIKRGNNLLTTGGTLKSIDGVAEALHPAADDTCTRSQSSPAFOPEIAN 226
QY 233 SEASTSGKADAVWSLGVALLFTMLAGHYPRODSHEPVLLFGKIRGAVALPAGLSAPARTY 292
Db 227 GUDTFSGEFVDIWSAGVTLNYITTTGUYPEEGDNLYLTFNIGKGSJAIPDCCGPLSDLL 286
QY 293 RCLTRREPARERLTANGILHPPMLROD-----EMLLAPTRSHLMEAAQVDP--DGIG 341
Db 287 KMLTYEPKAKRSIIOIRQHSWFRKKHPALPAEAPVPIPSBDTKDRKRSMTVVAYELDLHG 346
QY 342 LDEAREEEDGGRREVLY 357
Db 347 ADEDEDLFDIEDDITLY 362

```

```
Search completed: August 24, 2004, 18:54:08
Job time : 20 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 18:53:28 ; Search time 126 Seconds

(without alignments)
892.883 Million cell updates/sec

Title: US-10-070-337-5

Perfect score: 1892

Sequence: 1 MRATPLAAPAGSLSRKKRL.....GLGDDEAREEGREVVLYG 358

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1295152 seqs, 314255058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1888	99.8	358	US-09-799-875-8	Sequence 8, Appli
2	1888	99.8	358	US-10-649-156-8	Sequence 8, Appli
3	1884	99.6	398	US-10-425-114-37491	Sequence 37491, A
4	1884	99.6	404	US-10-425-114-53828	Sequence 53828, A
5	1859	98.3	360	US-10-408-765A-2189	Sequence 2189, A
6	1676	88.6	360	US-10-024-828-9	Sequence 1102, Ap
7	1102	58.2	263	US-09-925-301-1102	Sequence 304, App
8	678.5	35.9	269	US-10-221-278-304	Sequence 304, App
9	678.5	35.9	269	US-10-221-278-304	Sequence 304, App
10	656.5	34.7	290	US-10-221-278-680	Sequence 680, App
11	656.5	34.7	290	US-10-221-278-680	Sequence 680, App
12	575.5	30.4	205	US-10-228-263-2	Sequence 2, Appli
13	408	21.6	138	US-09-864-761-45767	Sequence 1367, Ap
14	354	18.7	153	US-09-925-301-1367	Sequence 1367, Ap
15	344	18.2	472	US-10-425-114-70164	Sequence 70164, A

16	341.5	18.0	443	US-10-437-963-179120	Sequence 179120,
17	318	16.8	461	US-10-437-963-174166	Sequence 174166,
18	305.5	16.1	373	US-10-026-021-174166	Sequence 4, Appli
19	305	16.1	477	US-10-437-963-122228	Sequence 122228,
20	298.5	15.8	404	US-10-437-963-188082	Sequence 188082,
21	296.5	15.7	778	US-10-423-543-11	Sequence 11, Appli
22	296.5	15.7	778	US-10-423-543-11	Sequence 11, Appli
23	296.5	15.7	778	US-10-423-543-11	Sequence 11, Appli
24	294.5	15.6	431	US-10-424-599-14642	Sequence 14642,
25	294	15.5	794	US-10-311-034-17	Sequence 17, Appli
26	293	15.5	437	US-10-424-599-189469	Sequence 189469,
27	293	15.5	460	US-10-425-114-55117	Sequence 55117, A
28	292	15.4	641	US-10-437-963-188700	Sequence 188700,
29	291.5	15.4	607	US-09-769-970-15	Sequence 15, Appli
30	291.5	15.4	607	US-10-108-580-2	Sequence 2, Appli
31	291.5	15.4	607	US-10-204-041-16	Sequence 16, Appli
32	291.5	15.4	607	US-10-620-052A-28	Sequence 28, Appli
33	290	15.3	434	US-10-425-114-49007	Sequence 49007, A
34	287.5	15.2	437	US-10-437-963-181140	Sequence 181140,
35	287.5	15.2	523	US-10-183-687-258	Sequence 258, App
36	287	15.2	511	US-10-437-963-125840	Sequence 125840,
37	286.5	15.1	459	US-10-425-114-66776	Sequence 66776, A
38	286.5	15.1	497	US-10-425-114-66776	Sequence 46452, A
39	286.5	15.1	754	US-10-479-532-1	Sequence 1, Appli
40	285.5	15.1	467	US-10-425-114-57045	Sequence 57045, A
41	284	15.0	1518	US-09-801-368-152	Sequence 152, App
42	284	15.0	1518	US-10-369-493-22243	Sequence 22243, A
43	284	15.0	1518	US-10-618-581-11	Sequence 11, Appli
44	283.5	15.0	448	US-10-424-599-236477	Sequence 236477,
45	283	15.0	504	US-10-183-687-407	Sequence 407, App

ALIGNMENTS

RESULT 1
US-09-799-875-8
; Sequence 8, Application US/09799875
; Patent No. US20020034780A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. US20020034780A1 Human Protein Kinases and Uses
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/09/799, 875
; CURRENT FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659,287
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-875-8

Query Match 99.8%; Score 1888; DB 9; Length 358;
Best Local Similarity 99.7%; Pred. No. 77e-159;
Matches 357; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRATPLAAPAGSLSRKKRLDNDLTERPVQRRASGPPRLPCLLPSPPTAPRRAT 60
DB 1 MRATPLAAPAGSLSRKKRLDNDLTERPVQRRASGPPRLPCLLPSPPTAPRRAT 60
QY AVATASRLGYVLLPEEGGRAYQALHCPGTCTCKVYVQALVLEPYARLPKRV 120
DB AVATASRLGYVLLPEEGGRAYQALHCPGTCTCKVYVQALVLEPYARLPKRV 120
QY 121 ARPEVLAVGQLVAFPTRTGDMHSIVTRRRIPEPEAAVLFQWATALAHCHQHGIVL 180

```
Db 121 ARPTEVLASTGLTLVAFETRTGDMHSLVRSRRIPEPEAAVLFRQWATLALHCHQGLVL 180
QY 181 RDLKLCRFVAFADREKRLVLENLDSCVLTGPDDSLMDKACPAVYGPETLSSRASYSKG 240
Db 181 RDLKLCRFVAFADREKRLVLENLDSCVLTGPDDSLMDKACPAVYGPETLSSRASYSKG 240
QY 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAVALPAGLSAPARCLVRCILRRP 300
Db 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAVALPAGLSAPARCLVRCILRRP 300
QY 301 AERLTATGILHLPWLRODPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358
Db 301 AERLTATGILHLPWLRODPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358
```

```
RESULT 2
US-10-649-156-8
; Sequence 8, Application US/10649156
; Publication No. US20040038346A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: No. US20040038346A1e1 Human Protein Kinases and Uses
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/10/649,156
; PRIOR FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/09/799,875
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659,287
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-649-156-8
```

```
Query Match 99.8%; Score 1888; DB 12; Length 358;
Best Local Similarity 99.7%; Pred. No. 7.7e-159;
Matches 357; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRATPLAAGSLSRKKRIELDNDLTERPVQKRAASGQPRLPCLPLSPPTADDRAT 60
Db 1 MRATPLAAGSLSRKKRIELDNDLTERPVQKRAASGQPRLPCLPLSPPTADDRAT 60
QY 61 AVATASRLGPPVYLLEBEEGGRAYOALHCPTEYTKVYPVOBALAVLEPYARLPKHAY 120
Db 61 AVATASRLGPPVYLLEBEEGGRAYOALHCPTEYTKVYPVOBALAVLEPYARLPKHAY 120
QY 121 ARPTEVLASTGLTLVAFETRTGDMHSLVRSRRIPEPEAAVLFRQWATLALHCHQGLVL 180
Db 121 ARPTEVLASTGLTLVAFETRTGDMHSLVRSRRIPEPEAAVLFRQWATLALHCHQGLVL 180
QY 181 RDLKLCRFVAFADREKRLVLENLDSCVLTGPDDSLMDKACPAVYGPETLSSRASYSKG 240
Db 181 RDLKLCRFVAFADREKRLVLENLDSCVLTGPDDSLMDKACPAVYGPETLSSRASYSKG 240
QY 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAVALPAGLSAPARCLVRCILRRP 300
Db 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAVALPAGLSAPARCLVRCILRRP 300
QY 301 AERLTATGILHLPWLRODPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358
Db 301 AERLTATGILHLPWLRODPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358
```

```
US-10-425-114-37491
; Sequence 37491, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37491
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Clone ID: LIB3101-194-B6_Flt.ped
US-10-425-114-37491
```

```
Query Match 99.6%; Score 1884; DB 12; Length 398;
Best Local Similarity 99.4%; Pred. No. 2e-158;
Matches 356; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRATPLAAGSLSRKKRIELDNDLTERPVQKRAASGQPRLPCLPLSPPTADDRAT 60
Db 41 MRATPLAAGSLSRKKRIELDNDLTERPVQKRAASGQPRLPCLPLSPPTADDRAT 100
QY 61 AVATASRLGPPVYLLEBEEGGRAYOALHCPTEYTKVYPVOBALAVLEPYARLPKHAY 120
Db 101 AVATASRLGPPVYLLEBEEGGRAYOALHCPTEYTKVYPVOBALAVLEPYARLPKHAY 160
QY 121 ARPTEVLASTGLTLVAFETRTGDMHSLVRSRRIPEPEAAVLFRQWATLALHCHQGLVL 180
Db 161 ARPTEVLASTGLTLVAFETRTGDMHSLVRSRRIPEPEAAVLFRQWATLALHCHQGLVL 220
QY 181 RDLKLCRFVAFADREKRLVLENLDSCVLTGPDDSLMDKACPAVYGPETLSSRASYSKG 240
Db 221 RDLKLCRFVAFADREKRLVLENLDSCVLTGPDDSLMDKACPAVYGPETLSSRASYSKG 280
QY 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAVALPAGLSAPARCLVRCILRRP 300
Db 281 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAVALPAGLSAPARCLVRCILRRP 340
QY 301 AERLTATGILHLPWLRODPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358
Db 341 AERLTATGILHLPWLRODPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 398
```

```
RESULT 4
US-10-425-114-53828
; Sequence 53828, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 53828
; LENGTH: 404
; TYPE: PRT
```


ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Clone ID: LIB082-008-C10_FLI pep
US-10-425-114-53828

Query Match 99.6%; Score 1884; DB 12; Length 404;
Best Local Similarity 99.4%; Pred. No. 2e-158;
Matches 356; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSRKRLLELDNDLTERPVOKRARGOPRLPCLPLSPPTADPDRAT 60
DB 47 MRATPLAAPAGSLSRKRLLELDNDLTERPVOKRARGOPRLPCLPLSPPTADPDRAT 106
QY 61 AVATASRLGPPVLLPEBEGGRAVQALHCPGTGTEYCKVYPVOEALAVLEPYARLPKHV 120
DB 107 AVATASRLGPPVLLPEBEGGRAVQALHCPGTGTEYCKVYPVOEALAVLEPYARLPKHV 166
QY 121 ARPEVLAGTQLLYAFPTRTGDMHSLVTRHRIPEPEAAVLFRQMATLALHCHQGLVL 180
DB 167 ARPEVLAGTQLLYAFPTRTGDMHSLVSRHRIPEPEAAVLFRQMATLALHCHQGLVL 226
QY 181 RDLKLCRFVADREKRLVLENDSCVLTGPDLSMDKACPAAYGPEILSSRASYSK 240
DB 227 RDLKLCRFVADREKRLVLENDSCVLTGPDLSMDKACPAAYGPEILSSRASYSK 286
QY 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVRLRRP 300
DB 287 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVRLRRP 346
QY 301 AERLTATGILLHPMLRQDMPPLAPTRSHLMEAAQVVDGLGDEAREEGDREVVLYG 358
DB 347 AERLTATGILLHPMLRQDMPPLAPTRSHLMEAAQVVDGLGDEAREEGDREVVLYG 404

RESULT 5

US-10-408-765A-2189
Sequence 2189, Application US/10408765A
Publication No. US20040101874A1
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Fahy, Bojin D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary M.
APPLICANT: Marnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
FILE REFERENCE: 660088-465
CURRENT APPLICATION NUMBER: US/10/408,765A
CURRENT FILING DATE: 2003-04-04
NUMBER OF SEQ ID NOS: 3077
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2189
LENGTH: 360
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: 193, 194, 195, 196, 197
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-408-765A-2189

Query Match 98.3%; Score 1859; DB 16; Length 360;
Best Local Similarity 98.3%; Pred. No. 2.9e-156;
Matches 354; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 1 MRATPLAAPAGSLSRKRLLELDNDLTERPVOKRARGOPRLPCLPLSPPTADPDRAT 60
DB 1 MRATPLAAPAGSLSRKRLLELDNDLTERPVOKRARGOPRLPCLPLSPPTADPDRAT 60
QY 61 AVATASRLGPPVLLPEBEGGRAVQALHCPGTGTEYCKVYPVOEALAVLEPYARLPKHV 120

DB 61 AVATASRLGPPVLLPEBEGGRAVQALHCPGTGTEYCKVYPVOEALAVLEPYARLPKHV 120
QY 121 ARPEVLAGTQLLYAFPTRTGDMHSLVTRHRIPEPEAAVLFRQMATLALHCHQGLVL 180
DB 121 ARPEVLAGTQLLYAFPTRTGDMHSLVSRHRIPEPEAAVLFRQMATLALHCHQGLVL 180
QY 181 RDLKLCRFVADREKRLVLENDSCVLTGPDLSMDKACPAAYGPEILSSRASYS 238
DB 181 RDLKLCRFVADREKRLVLENDSCVLTGPDLSMDKACPAAYGPEILSSRASYS 240
QY 239 GKADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVRLRR 298
DB 241 GKADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVRLRR 300
QY 299 EPARLTATGILLHPMLRQDMPPLAPTRSHLMEAAQVVDGLGDEAREEGDREVVLYG 358
DB 301 EPARLTATGILLHPMLRQDMPPLAPTRSHLMEAAQVVDGLGDEAREEGDREVVLYG 360

RESULT 6

US-10-024-828-9
Sequence 9, Application US/10024828
Publication No. US20030036051A1
GENERAL INFORMATION:
APPLICANT: Virca, Duke
APPLICANT: Bird, Timothy A.
APPLICANT: Anderson, Dirk M.
APPLICANT: Marken, John S.
TITLE OF INVENTION: Human CDNA's Encoding Polypeptides Having Kinase
TITLE OF INVENTION: Functions
FILE REFERENCE: 2877-US
CURRENT APPLICATION NUMBER: US/10/024,828
CURRENT FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: US/09/509,902A
PRIOR FILING DATE: 1999-08-03
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 360
TYPE: PRT
ORGANISM: Homo sapiens
US-10-024-828-9

Query Match 88.6%; Score 1676; DB 14; Length 360;
Best Local Similarity 99.7%; Pred. No. 4.9e-140;
Matches 317; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSRKRLLELDNDLTERPVOKRARGOPRLPCLPLSPPTADPDRAT 60
DB 43 MRATPLAAPAGSLSRKRLLELDNDLTERPVOKRARGOPRLPCLPLSPPTADPDRAT 102
QY 61 AVATASRLGPPVLLPEBEGGRAVQALHCPGTGTEYCKVYPVOEALAVLEPYARLPKHV 120
DB 103 AVATASRLGPPVLLPEBEGGRAVQALHCPGTGTEYCKVYPVOEALAVLEPYARLPKHV 162
QY 121 ARPEVLAGTQLLYAFPTRTGDMHSLVTRHRIPEPEAAVLFRQMATLALHCHQGLVL 180
DB 163 ARPEVLAGTQLLYAFPTRTGDMHSLVSRHRIPEPEAAVLFRQMATLALHCHQGLVL 222
QY 181 RDLKLCRFVADREKRLVLENDSCVLTGPDLSMDKACPAAYGPEILSSRASYSK 240
DB 223 RDLKLCRFVADREKRLVLENDSCVLTGPDLSMDKACPAAYGPEILSSRASYSK 282
QY 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVRLRRP 300
DB 283 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVRLRRP 342
QY 301 AERLTATGILLHPMLRQD 318
DB 343 AERLTATGILLHPMLRQD 360

RESULT 7

US-09-925-301-1102
; Sequence 1102, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1102
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1102

Query Match 58.2%; Score 1102; DB 9; Length 233;
Best Local Similarity 100.0%; Pred. No. 2,2e-89;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 RHRIPEEAALFRONATLAHQHGLVRLDKLCRFVADREERKLVLENEDSCVLT 210
DB 26 RHRIPEEAALFRONATLAHQHGLVRLDKLCRFVADREERKLVLENEDSCVLT 85
QY 211 GPDDSLMDKACPAVYGPETLSRASYSKGAADVWSLGVALLFTMLAGHYFPQSEPVLLF 270
DB 86 GPDDSLMDKACPAVYGPETLSRASYSKGAADVWSLGVALLFTMLAGHYFPQSEPVLLF 145
QY 271 GKIRGAYALPAGLSAPARCLVRCILRREPARELTATGILLHFWLQDPPMLAPTRSHLM 330
DB 146 GKIRGAYALPAGLSAPARCLVRCILRREPARELTATGILLHFWLQDPPMLAPTRSHLM 205
QY 331 EAAQVVPDGLGDEAREEEDREVVLVYG 358
DB 206 EAAQVVPDGLGDEAREEEDREVVLVYG 233

RESULT 8
US-10-221-278-304
; Sequence 304, Application US/10221278
; Publication No. US20040034208A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/221,278
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 304
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-221-278-304

Query Match 35.9%; Score 678.5; DB 12; Length 269;
Best Local Similarity 51.4%; Pred. No. 9.4e-52;

Matches 132; Conservative 39; Mismatches 79; Indels 7; Gaps 2;
QY 88 CP-----TGTETXCKYYPVQDALVLEPYARLPKHKVAPRTVLAGTQLLYAFRTTH 141
DB 2 CPGRCASTLGRVYCKVFPIKHVQDKIRPYIQLPESHNITGIVILLGETATVYFFEKDF 61
QY 142 GDMHSIVTRRHRIPEEAALFRONATLAHQHGLVRLDKLCRFVADREERKLVLE 201
DB 62 GDMHSIVTRRHRIPEEAALFRONATLAHQHGLVRLDKLCRFVADREERKLVLE 121
QY 202 NLEDSCVLTGPDSDLMWKACPAVYGPETLSRASYSKGAADVWSLGVALLFTMLAGHYFP 261
DB 122 SLSDTHIMKCEDDALSKHGCAPVYSEIINTGTISGKADVWSLGVALLFTMLAGHYFP 181
QY 262 QDSEPVLLFGKIRGAYALPAGLSAPARCLVRCILRREPARELTATGILLHFWLQDPP 321
DB 182 HSDSPSALFSGKIRRGQCFIPEHISPKARCLIRSLRREPARELTATGILLHFWLQDPP 241
QY 322 LAPTRSHLMFAQVVPD 338
DB 242 -GYTDSBITGSDQIVPE 257

RESULT 9
US-10-291-172-304
; Sequence 304, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 304
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-172-304

Query Match 35.9%; Score 678.5; DB 15; Length 269;
Best Local Similarity 51.4%; Pred. No. 9.4e-52;
Matches 132; Conservative 39; Mismatches 79; Indels 7; Gaps 2;

QY 88 CP-----TGTETXCKYYPVQDALVLEPYARLPKHKVAPRTVLAGTQLLYAFRTTH 141
DB 2 CPGRCASTLGRVYCKVFPIKHVQDKIRPYIQLPESHNITGIVILLGETATVYFFEKDF 61
QY 142 GDMHSIVTRRHRIPEEAALFRONATLAHQHGLVRLDKLCRFVADREERKLVLE 201
DB 62 GDMHSIVTRRHRIPEEAALFRONATLAHQHGLVRLDKLCRFVADREERKLVLE 121
QY 202 NLEDSCVLTGPDSDLMWKACPAVYGPETLSRASYSKGAADVWSLGVALLFTMLAGHYFP 261
DB 122 SLSDTHIMKCEDDALSKHGCAPVYSEIINTGTISGKADVWSLGVALLFTMLAGHYFP 181
QY 262 QDSEPVLLFGKIRGAYALPAGLSAPARCLVRCILRREPARELTATGILLHFWLQDPP 321
DB 182 HSDSPSALFSGKIRRGQCFIPEHISPKARCLIRSLRREPARELTATGILLHFWLQDPP 241
QY 322 LAPTRSHLMFAQVVPD 338

QY 264 SEPVLLFGKIRRGAYALPAGLSAPARCLVRLRREPARELTATGILHPWLRODPMLA 323
Db 121 SDPSALFSKIRRGQCFEPHISPARCLIRSLRREPSERLTAEILHPFESVLEP-G 179
QY 324 PTRSHLWEAQQVVD 338
Db 180 YIDSEICTSDQVPE 194

RESULT 13
US-09-864-761-45767
Sequence 45767, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
NUMBER OF SEQ ID NOS: 2001-01-29
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 45767
LENGTH: 138
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC009486.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.93
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.93
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96

OTHER INFORMATION: SWISSPROT HIT: O74536, EVALUATE 1.00e-18
OTHER INFORMATION: EST_HUMAN HIT: BB97149.1, EVALUATE 2.00e-71
US-09-864-761-45767
Query Match 21.6%; Score 408; DB 9; Length 138;
Best Local Similarity 63.1%; Pred. No. 3.8e-28;
Matches 77; Conservative 15; Mismatches 30; Indels 0; Gaps 0;
QY 197 KLVLENLEDSCVLTGPDSDIMDKHACPAVYGPETLSSRASYSGRADWMSIGVALFTMLA 256
Db 2 RVLKESIEDAVILRGDDSDSKHGCAPVYSPETLNTSSGSGRADWMSIGVALFTMLV 61
QY 257 GHYFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRLRREPARELTATGILHPWL 316
Db 62 GRYPFHDIKPSLSFKIRRGQFNIPELTSPKACLIIRSLRREPSERLTQELIDHPWFS 121
QY 317 QD 318
Db 122 TD 123

RESULT 14
US-09-925-301-1367
Sequence 1367, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1367
LENGTH: 153
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (136)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (138)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (141)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (142)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (143)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (152)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1367
Query Match 18.7%; Score 354; DB 9; Length 153;
Best Local Similarity 93.3%; Pred. No. 2.7e-23;
Matches 70; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 MRATPLAAPAGSUSRRKRLLELDNLDTERPVOKARSGPQRRLPPCLPLSPPTAPDRAT 60
Db 74 MRATPLAAPAGSUSRRKRLLELDNLDTERPVOKARSGPQRRLPPCLPLSPPTAPDRAT 133
QY 61 AVATASRLGPVYLE 75
Db 134 AVXTXSRXXXYVILE 148

RESULT 15

US-10-425-114-70164
 ; Sequence 70164, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 70164
 ; LENGTH: 472
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: UC-ZMFLB73099C12_F11.pep
 US-10-425-114-70164

Query Match 18.2%; Score 344; DB 12; Length 472;

Best Local Similarity 33.3%; Pred. No. 9e-22; Indels 28; Gaps 10;

Matches 98; Conservative 44; Mismatches 124;

```

QY 51 SP-PTAPPRAVATASR--LGPVYL--LPEEGGRAVQALHCPTEYTCVKYVQEA 104
Db 3 SPSPQPD---MVSQARGLGAIELGRTLGEGNFQKVKQARHRSQGFAYK--IMER 56
QY 105 IAVLEPYA-----RLPPKHVARFTEVLAGTQLLYAFTRTHGD--MSIVRTR 151
Db 57 ARVLQGRADQICREIATLKLAAHPVAVLHHEVAASKTKIYVLELVNNGELIDRIASE 116
QY 152 HRIPEPAVLFROMATALAHQHGLVLRDLKLCRFVADRERK-KLVLENLEDCVLT 210
Db 117 GKLPQEARRLFOQLVDGVSCHKEGVCHRDLT--ENVLVRKNIKISDFGISALPQHL 175
QY 211 GDDDSIMDMHACPAYVGPETLSSRASYSQKADVMSIGVALFTMLAGHYFPQDSEPVLLF 270
Db 176 GNDGLHTTCGSPNTAFEVLONR-GYDGLSDIWSGCVILYVMLVGHLPFDDRNIVLY 234
QY 271 GKIRGAYALPAGISAPARCLVRCLLRREPAPERLTATGILLHPMLRQDPMPLAP 324
Db 235 QKIFKGDAGQFPEWLSPGARNTLRLRILEPDAERIAMAEIKAHMPWFQEHVVFVLP 288

```

Search completed: August 24, 2004, 19:04:28

Job time : 129 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 18:50:17 ; Search time 19 seconds
(without alignments)
1812.451 Million cell updates/sec

Title: US-10-070-337-5

Perfect score: 1892
Sequence: 1 MRATPLAAPAGSLSRKKRL.....GLGLDEAREEGDREVLYXG 358

Scoring table: BIOSDM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	299.5	15.8	631	2	A57286
2	298	15.8	887	2	T20941
3	291.5	15.4	651	2	S52244
4	290	15.3	602	2	S72513
5	284	15.0	1518	2	S37928
6	283	15.0	504	2	T10449
7	283	15.0	512	2	T52633
8	280	14.8	512	1	UC1446
9	279	14.7	339	2	S56719
10	276.5	14.6	502	2	T02306
11	276	14.6	441	2	E85362
12	276	14.6	511	1	A56009
13	274	14.5	431	2	T02496
14	273.5	14.5	633	1	A26030
15	273	14.4	513	1	S60304
16	272	14.4	504	2	T07415
17	271.5	14.3	426	2	C71408
18	270	14.3	576	2	T41587
19	269.5	14.2	480	2	A86477
20	269	14.2	512	2	T07788
21	268.5	14.2	421	2	E96522
22	268.5	14.2	533	1	A34366
23	267.5	14.1	593	1	KIZPMN
24	266	14.1	472	2	B90100
25	266	14.1	542	1	A45025
26	264	14.0	622	1	S44859
27	263	13.9	542	1	A26464
28	263	13.9	589	2	S68470
29	262.5	13.9	421	2	T48202

30	262	13.8	1142	2	S59359	G1N4 protein - yea
31	261	13.8	442	2	T48203	hypothetical prote
32	260.5	13.8	513	1	S60303	serine/threonine-s
33	260.5	13.8	552	1	S51025	hydoxymethylglut
34	259.5	13.7	473	1	S59941	serine/threonine-s
35	259.5	13.7	552	1	A53621	hydoxymethylglut
36	259.5	13.7	891	2	T40503	protein kinase kin
37	259.5	13.7	891	2	A38903	protein kinase 1 -
38	258	13.6	350	2	T06107	hypothetical prote
39	258	13.6	726	2	T33998	hypothetical serine/th
40	257.5	13.6	440	2	T14736	calcium-dependent
41	257.5	13.6	492	1	T03271	calcium-dependent
42	257	13.6	469	2	B84644	probable protein k
43	257	13.6	502	1	A41361	serine/threonine-s
44	255.5	13.5	445	2	T50802	serine/threonine p
45	255	13.5	610	1	A49082	calcium-dependent

ALIGNMENTS

RESULT 1

A57286
probable serine/threonine protein kinase (EC 2.7.1.-) fnk - mouse
C:Species: Mus musculus (house mouse)
C:Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 10-Sep-1997
C:Accession: A57286
R:Donohue, P.D.; Alberts, G.F.; Guo, Y.; Winkles, J.A.
J. Biol. Chem. 270, 10351-10357, 1995
A:Title: Identification by targeted differential display of an immediate early gene enc
A:Reference number: A57286; MUID:95247749; PMID:7730342
A:Accession: A57286
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-631 <DON>
A:Cross-references: GB:U21392; GB:U22434
C:Superfamily: unassigned Ser/thr or Tyr-specific protein kinases; protein kinase homol
C:Keywords: ATP; phosphotransferase
F:61-315/Domain: protein kinase homology <KIN>

Query Match 15.8%; Score 299.5; DB 2; Length 631;
Best Local Similarity 28.3%; Pred. No. 3,8e-13;
Matches 98; Conservative 48; Mismatches 133; Indels 67; Gaps 14;

QY	39	PPPRPCLLPSPPTAPRATAVATAS-RLGPVYLPE-----	77
DB	10	PPR-PPRAVPSPAPAPG--PPNAPSREPEVLACGRADPPGRLLTDLDSGRYTK 65	
QY	78	-----EGG--RAYQALHCPTGTEYCKVYPVOEALAVLEPYARLP-----PKKA 119	
DB	66	GRLLGKGPRACRYEARDTDSGLAVAKVLPQR--VAKPHQREKILNIEELHRDLQHRH 122	
QY	120	VAPPEVLATGOLLYAFTR--THGDMHSLVTRRHRPEPEAVLPROMATLAHQHGH 178	
DB	123	YRFSHFEDADNIVFELICSRKSLAHMKARHTLEBVRYYQLISGLKYDHORG 182	
QY	179	VLRDLKCRFPVPRDRERKLVLENLSDCVLTGPDSDMDKAC--PAYGPEILSSRAS 236	
DB	183	LRHDLKGNFFITD--NMELKVGDFGLARLEPPEQR--KKTICGTPNYAPVLLRQG- 237	
QY	237	YSGKADVMSLGVAFETMLAGHPFODSEFVLLFGKIRGAYALPAGLSAPARCLVACL 296	
DB	238	-HPEADVMSLGCVMITLLCGSPFETADKEHYRCIKQVHTLPASLSIPARQLLAITL 296	
QY	297	RREPARLITATGILLHPWLQ---DMPILAPTRSHMEAAQVVPD 338	
DB	297	RASPRDRPSLEQLIRHDFFTKGYTPDRLPVS-----SCVTVPD 334	

RESULT 2
T20941
hypothetical protein F15A2.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C/Accession: T20941
R:Gregory, J.
Submitted to the EMBL Data Library, March 1996
A/Reference number: Z19349
A/Accession: T20941
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-887 <WILL>
A/Cross-references: EMBL:Z70207, PIDN:CAA94127.1, GSPDB:GN00028, CESP:FL5A2.6
A/Experimental source: clone FL5A2
C/Genetics:
A/Gene: CESP:FL5A2.6
A/Map position: X
A/Intons: 32/1; 63/3; 92/2; 139/2; 189/3; 328/2; 448/2; 516/3; 604/2; 684/3; 735/2; 777

Query Match 15.8%; Score 298; DB 2; Length 887;
Best Local Similarity 25.9%; Pred. No. 7e-13;
Matches 95; Conservative 51; Mismatches 133; Indels 88; Gaps 14;

QY 62 VATASTRIGYVL--LEPEEGRAYQALHCPTEGYCKV-----YVQELAV 107
DB 11 VAAQAYCGPYKLTGKQGTGLVKGTGTCITGRKAIVKNEKLSVLYQVEREIAL 70
QY 108 LEFYALRPHKHVARPTVLAQTOLYAFTR--THGDMSLVTRRRIPEPEAAVLFROV 166
DB 71 M-----KLIEHPVHLHYDYENKRYLYLLLEHVSGLFEDYLVKRGRLMKARKFROQ 126
QY 167 ATALAHCHQGLVLDLCLCFEFADEKREK-----LVEN--LESCVITGDDSL 216
DB 127 ISALDFCHAHNIQRLKXENLLDERNNIKAVDFGMSIQVSGSMLETSQ----- 177
QY 217 WDKHACPAVGPBILSSRASYSGKADVSLGVALFTMLAGHPFODSEPVLLFGKIRG 276
DB 178 ----GSPHYACEVIRGE--KYDGRKADVMSGCVLLVALVGLPFPDDMLRLNLEKVKRG 232
QY 277 AYALPAGASAPARCIVRCILRREPARLTATGILLHPLT-----RQDP--MPLAP--TRSH 328
DB 233 VHIHFVFPADVQSLIRAMIEVDPGKRYSLADVFKHPWVGTTKADPELELPMQVQYQH 292
QY 339 IWEAAQV--PDGLG-----IDENAE--EGGD 351
DB 293 VIFGEDSIDPVLRLMNCIGCFKDKOKILNELSFNANTKRVYFLLLDRKRREPDAQDD 352
QY 352 REVLYYG 358
DB 353 TEIVLRG 359

RESULT 3

S52244
p69g3 protein - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 07-May-1995 #sequence_revision 03-Aug-1995 #text_change 24-May-2001
C/Accession: S52244
R/Rochi, C.; le Guellec, R.; Paris, J.; Couturier, A.; Philippe, M.
Submitted to the EMBL Data Library, October 1992
A/Description: E93, selected by differential screening encodes a new Xenopus protein kin
A/Reference number: S52243
A/Accession: S52244
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-651 <ROG>
A/Cross-references: EMBL:Z17205, MID:9609283, PIDN:CAA78913.1, PID:9609284
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: ATP
F:11-265/Domain: protein kinase homology <KIN>
F:19-27/Region: protein kinase ATP-binding motif

Query Match 15.4%; Score 291.5; DB 2; Length 651;
Best Local Similarity 29.9%; Pred. No. 1.4e-12;
Matches 73; Conservative 47; Mismatches 103; Indels 21; Gaps 7;

QY 85 ALHCPTEGYCKXYVPOEALAVLEPARLP-----PHKGVAPTEVLAGTOLYAF 137
DB 30 ASHLITGEKVAIKIMD--KESLGDLPKVTEDIMKMLSHQVCHLVITPKKIFVYL 88
QY 138 TR--THGDMSLVTRRRIPEPEAVLRQATALAHCHQGLVLRDLCLCFVADREK 196
DB 89 EYCPGSELFDYITAKDLTEEARVFFROIVSAVAVYHSQYARLDKPEMLIDEDQNL 148
QY 197 KLVLENLDESCVLTPDPSLMDKH-----ACPAVGPBILSSRASYSGKADVSLGVAL 251
DB 149 KLI-----DFGLCAKPKGGL--DYHLMTCGSPAYABELLQGRK--YIGSERDITMSWGLM 201
QY 252 FTMLAGHPPODSEPVLLFKIRRGAYADPAGSAPARCIVRCILRREPARLTATGILL 311
DB 202 YALMCGLFPDDDNVWLVIKIMRGKYEIPKWLSPGVSLLSQMVQDPKRRITVHILN 261
QY 312 HPWL 315
DB 262 HPWL 265

RESULT 4

S72513
FOG2 protein - yeast (Kluyveromyces marxianus var. lactis)
C/Species: Kluyveromyces marxianus var. lactis, Candida sphaerica
C/Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 07-May-1999
C/Accession: S72513
R/Boffin, P.; Ficarella, A.; Domini, C.; Lodi, T.; Pngli, P.P.; Ferrero, I.
Curr. Genet. 29, 316-326, 1996
A/Title: FOG1 and FOG2 genes, required for the transcriptional activation of glucose-reg
A/Reference number: S72513; MID:96171514; PMID:8598052
A/Accession: S72513
A/Status: not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-602 <GOF>
A/Note: the source is designated as Kluyveromyces lactis
C/Genetics:
A/Gene: FOG2
C/Function:
A/Description: probably involved in the regulation of glucose-repressible gene expressio
A/Superfamily: AMP-activated protein kinase; protein kinase homology
F:33-286/Domain: protein kinase homology <KIN>

Query Match 15.3%; Score 290; DB 2; Length 602;
Best Local Similarity 28.5%; Pred. No. 1.6e-12;
Matches 88; Conservative 51; Mismatches 138; Indels 32; Gaps 10;

QY 63 ATASRLIGYVLEP--EES--GRAYQALHCPTEGYCKVYVQELAVLE-----P 110
DB 27 AOGGHIGKQIITKTLGSGSFQKVLAVHISTGQVAKLIN--KVVLAKSDMQRIEREIS 85
QY 111 YARLPFHKHVARPTVLAQTOLYAFTRTHGDMSLVTRRRIPEPEAAVLFROV 170
DB 86 YLRRLRPHIITKLDVILKSDQEIIMVIEYANGLFDYIVQDKMBQBARFPQOIIISAV 145
QY 171 AHCHQGLVLRDLCLCFVADREKRLVLENLDESCVLTPDPSLMDKHAC--PAYVGP 228
DB 146 DYCHRHICIVRDLPEMLIDENHINVKIADFG--SNIMT--DGNFLKISCGSPNNAP 200
QY 229 EILSSRASYSGKADVSLGVALFTMLAGHPFODSEPVLLFGKIRGAYALPAGSAPA 288
DB 201 EVISGKL--YAGPEVDWSSGVLLYVMLCRRLPFODESIPVLFKNIISGVYTIIPNPLSGA 259
QY 289 RCLVRCILRREPARLTATGILLHPLRQD--PMPDAPTRSHLWEEAAQVVDGLDRE 347
DB 260 ASLTKKMLVNPVNPRIIVHIMQDEMPKULLPVLIVASTHQNLS-----ESKT 309
QY 348 EEGDREYVL 356
DB 310 EDDGPSVPL 318

RESULT 5


```

Db      32 GKVIAEIVLVGHKVAIKILNRKIKOMMEMEKURREIKIL-----RLFMHHIIRLOEVI 87
QY      128 AGTQLVAFPTRT-HGDMHSLVTRRHRIPEEPAVLPRQATATALAHQHGVLRIIDK- 184
Db      88 ETTSDIIVMEVYVSGELFDYIVVEKGRLOEDEANRFQOIISGYEYCHRRNVVHRDJKPE 147
QY      185 -----LCRFVFADERRERKVLLEN--LEDSCVLTPGDDSLMDKACAPAYGPELTSSRAS 236
Db      148 NLLDSRCNKVIADFGLSNMVRDGHFLKTSC-----GSPNYAAPEVTSGL- 193
QY      237 YSGKAADVMSLGVALLFTMLAGHYFPDSEVLLFGKIRGAVALPAGISAPARCLVRCCL 296
Db      194 YAGEVDVMSCGVILVALLCGTLFPDDENIPNLFKTKIGGITYLTPSHLSEARDLIRML 253
QY      297 RREPARELTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVPDGLDEARE--EEGREGV 354
Db      254 IVDPMKRVITPEIRQHPWF-----QTHLPRTYLAVSPPT-VEQAKKINEEIVQEV 302
QY      355 VLYG 358
Db      303 VNYG 306

```

RESULT 8

JCI446

serine/threonine-specific protein kinase (EC 2.7.1.-) AK21 - Arabidopsis thaliana
N/Alternate names: protein kinase SNF1 homolog
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #ext_change 11-Jun-1999
C/Accession: JCI446; S58266; S66334

R/RefSeq: L. Thomas, M.; Blanch, M.; Halford, N.G.; Kreis, M.
Gene 120, 249-254, 1992

A/Title: Structure and expression of a gene from Arabidopsis thaliana encoding a protein

A/Reference number: JCI446; MUID:93013041; PMID:1339373

A/Accession: JCI446

A/Molecule type: DNA

A/Residues: 1-512 <LEG>

A/Cross-references: GB:M93023; NID:g166599; PIDN:AAA2736.1; PID:g166600

R/Thummler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.
submitted to the EMBL Data Library, May 1995

A/Description: Differential accumulation of the transcripts of 22 novel protein kinase

A/Reference number: S58256

A/Accession: S58266

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 144-198 <THU>

A/Cross-references: EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PID:g928910

R/Thummler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.
Plant Mol. Biol. 29, 551-565, 1995

A/Title: Differential accumulation of the transcripts of 22 novel protein kinase genes

A/Reference number: S66334; MUID:96123233; PMID:8534852

A/Accession: S66334

A/Molecule type: DNA

A/Residues: 144-198 <TH2>

A/Cross-references: EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PID:g928910

C/Comment: This enzyme plays an important role in a signal transduction cascade regulat

C/Genetics:

A/Genes: AK10; AK21

A/Introns: 64/1; 125/3; 186/3; 230/3; 292/3; 322/3; 350/3; 396/3; 475/3

C/Function:

A/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin

C/Superfamily: Am-activated protein kinase; protein kinase homology

C/Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase

F/17-271/Domain: protein kinase homology <KIN>

F/25-33/Region: protein kinase ATP-binding motif

F/48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted

F/147,151/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 14.8%; Score 280; DB 1; Length 512;

Best Local Similarity 28.7%; Pred. No. 6.5e-12;

Matches 87; Conservative 46; Mismatches 118; Indels 52; Gaps 9;

```

QY      80 GRAVQALHCPTEYTCVY-----PYQALAVDEPYARLPKHVAPTEVL 127
Db      31 GRVIAEHALTGKVAIKILNRKIKOMMEMEKURREIKIL-----RLFMHHIIRLOEVI 86
QY      128 AGTQLVAFPTRT-HGDMHSLVTRRHRIPEEPAVLPRQATATALAHQHGVLRIIDK- 184
Db      87 ETTSDIIVMEVYVSGELFDYIVVEKGRLOEDEANRFQOIISGYEYCHRRNVVHRDJKPE 146
QY      185 -----LCRFVFADERRERKVLLEN--LEDSCVLTPGDDSLMDKACAPAYGPELTSSRAS 236
Db      147 NLLDSRCNKVIADFGLSNMVRDGHFLKTSC-----GSPNYAAPEVTSGL- 192
QY      237 YSGKAADVMSLGVALLFTMLAGHYFPDSEVLLFGKIRGAVALPAGISAPARCLVRCCL 296
Db      193 YAGEVDVMSCGVILVALLCGTLFPDDENIPNLFKTKIGGITYLTPSHLSEARDLIRML 252
QY      297 RREPARELTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVPDGLDEAREEGBREV 355
Db      253 VVDPMKRVITPEIRQHPWF-----QAHLPRTYLAVSPPTVQOAKKIDEIIOEVI 302
QY      356 LYG 358
Db      303 NMG 305

```

RESULT 9

S56719

serine/threonine-specific protein kinase SPK-1 (EC 2.7.1.-) - soybean
C/Species: Glycine max (soybean)
C/Date: 28-Oct-1995 #sequence revision 03-Nov-1995 #ext_change 18-Jun-1999
C/Accession: S56719

R/RefSeq: Yoon, H.W.; Jeong, Y.H.; Bahk, J.D.; Hong, J.C.; Cho, M.J.
submitted to the EMBL Data Library, January 1993

A/Description: Cloning of a novel protein serine/threonine kinase cDNA from soybean.

A/Reference number: S56719

A/Accession: S56719

A/Molecule type: mRNA

A/Residues: 1-339 <SHI>

A/Cross-references: EMBL:L01453; NID:g169990; PIDN:AAA33979.1; PID:g169991

C/Superfamily: kinase-related transforming protein; protein kinase homology

C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F/3-261/Domain: protein kinase homology <KIN>

F/11-19/Region: protein kinase ATP-binding motif

Query Match 14.7%; Score 279; DB 2; Length 339;

Best Local Similarity 29.3%; Pred. No. 4.8e-12;

Matches 86; Conservative 34; Mismatches 86; Indels 88; Gaps 11;

```

QY      117 HKHVARPTREVLAGTQLVAFPTRT-----GDMHSLVTRRHRIPEEPAVLPRQ 166
Db      58 HPNIIRFEV-----FLPFTHALVLEYYAAGELFERICNAGRUSEDEARFFQOL 108
QY      167 ATALAHQHGVLRIIDK-----CRFVFADERRERKVLLENEDSCVLTG 211
Db      109 ISGVSYCHSMGICHRDLKENTLLDGNPARLTKCDPEFS--KSAIHSQPKSTVGT- 163
QY      212 PDDSLMDKACAPAYGPELTSSRASYSKGAADVMSLGVALLFTMLAGHYFPDSEVLLF- 270
Db      164 -----PAYIADEVL-SRKEYDGKADVMSCGVTLYVMLVGAVPEPDPEPKNFR 211
QY      271 ---GKIRGAVADP--AGLSAPARCLVRCULRREPARELTATGILLHPMLRQD-PMPL- 322
Db      212 KSIGRIMSVQYALPDYVRSKCHSLSCIFVANPARISISELKHHPFKNLPREITE 271
QY      323 -----APTRSHLMEAAQV-VPDGLDEAREE 348
Db      272 PERRGVVDHKTQPSQSVERTMRIQEARIKIHGEQAGTSTDAVAGDEANEE 325

```

RESULT 10

T02306

probable protein kinase [imported] - Arabidopsis thaliana
N/Alternate names: hypothetical protein F13p17.2

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C:Accession: J02306; D84753
C:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, July 1998
A:Description: Arabidopsis thaliana chromosome II BAC F13p17 genomic sequence.
A:Reference number: Z14657
A:Accession: T02306
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-502 <ROU>
A:Cross-references: EMBL:AC004481; NID:G3337347; PIDN:AAC27394.1; PID:G3337349
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Varkken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eskin, U.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: D84753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-502 <STO>
A:Cross-references: GB:AE002093; NID:G3337349; PIDN:AAC27394.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g34180; F13P17.2
A:Map position: 2
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
F:55-311/Domain: protein kinase homology <KIN>

```

Query Match      14.6%; Score 276.5; DB 2; Length 502;
Best Local Similarity 26.0%; Pred. No. 1.1e-11;
Matches 89; Conservative 58; Mismatches 13; Indels 61; Gaps 12;

Qy      3 APLIAP-----AGSLRKKRLLELDNDLTERPVYKRRASRGQPRLLPLSPT 54
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      6 STLPLAPGPIPIQFMAGLAR--IVTKNTNETSPSPRS-----PRP----- 48

Qy      55 ADRPATAVATASRLGPVYLLPEPEGRAVYQALHCTGTBYTKVY-----VQ 102
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      49 ---GGSILMKVEIGK--LLHGSGFAKYVLLRNINSHGSDVAIKVDKEIKVSGLAGHK 103

Qy      103 ELALVLEPVARLPKHVARPTEVLAQQLYAFTQH--GMSILVTRHRIPPEAV 161
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      104 REISIL---RRVHPYIVHLEVMATKTKIYIWEYVARGELVNTV--ARGLREGTARR 158

Qy      162 LPRQMATALAHCHQGLVRLDKLCRFVADRERKKVLLENIDSCVLTGPDSDMDKHA 221
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      159 YFOOLISSVAFCHSRGVYHRDKLENLLDDKGNKV-----SPGSLVSEQLKQEGT 212

Qy      222 C-----PAYVPELLSRASYSGKAADVWSIGVLLFTMLAGHYRQDSVPVLLPEKIR 275
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      213 CQTFCGTPAYIAPVLT--TRKGYEGAKADIWSGVLLFVIMAGSLFPDDKNILVMTKTIYK 271

Qy      276 GAYALPAGLSAPARCLVCLRRPEAPERLTATGILLHPWLQ 317
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      272 GQFCRKPMSPEELARLVTRMDLTNDPTRIITPEIMKHRRFKK 313

RESULT 11
E85362
hypothetical protein AT4g30960 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #ext_change 02-Mar-2001
C/Accession: E85362
R/anonymous: The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A/Reference number: A85001; MUID:20083488; PMID:10617198
A/Accession: E85362
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-441 <STO>

```

A:Cross-references:	GB:NC_001268; NID:g7269998; PIDN:CAB79814.1; GSPDB:GN00140
C:Genetics:	
A:Gene:	AT4g30960
A:Map position:	4
C:Superfamily:	unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
Query Match	14.6%; Score 276; DB 2; Length 441;
Best Local Similarity	29.4%; Pred. No. 1e-11;
Matches	77; Conservative 39; Mismatches 125; Indels 22; Gaps 4;
QY	81 RAYQALHCPGTETCYCKY-----PVCBALAVLEPYARLPKHKVARPTETVA 128
DB	37 KYVHARNITQGTGKSYAMKRVGKEKRVKVGMDQIKRELSVM---RMVGHPIINVELHEWYA 92
QY	129 GTOLLIAFTTTHGDMHSLVTRHRIPEPEAAVLFRQMATALAHCHQGVLRDLKCRF 188
DB	93 SKSKTIYFAMEIYRGSELPAKYAKGRLREDVARYVFQOLISAVDFCHSRGCVYHRDLKPENTL 152
QY	189 VFADPERKKLVLENIEDSCVLTPDDSLMDHACAPAYGPEITLSRASYSKGAADWSIG 248
DB	153 LIDEGNKLVKVDPGISAFPTETHLKODGLDHTTCGIPAYVAPEVITLKK-gydgkkaadlmscg 211
QY	249 VALFTMLAGHYPEQDSEPVLLFGKIRRGAYVLPAGLSAPARCLVCLRRPEARLITATG 308
DB	212 VILFPLLLGLGYLPQDDNLVNMVTRKIYRGDFKCPGWLSSDARLVTKLDDPENTRITLTK 271
QY	309 ILIHFWLRQDPM-----PLAPT 325
DB	272 VMDSEFWFKQATRSRKNPEVAAT 293

```

RESULT 12
A56009
serine/threonine-specific protein kinase (EC 2.7.1.-) NPK5 - common tobacco
C|Species: Nicotiana tabacum (common tobacco)
C|Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 16-Jun-2000
C|Accession: A56009
R:Murataka, T.; Banno, H.; Machida, Y.
Mol. Cell. Biol. 14, 2958-2965, 1994
A|Title: Characterization of tobacco protein kinase NPK5, a homolog of Saccharomyces cerevisiae
tase of Saccharomyces cerevisiae
A|Reference number: A56009; MUID:94217693; PMID:8164654
A|Accession: A56009
A|Status: preliminary
A|Molecule type: mRNA
A|Residues: 1-511 <MUR>
A|Cross-references: GB:D26602; NID:3496384; PIDN:BA05649.1; PID:3496385
C|Function:
A|Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine-phosphate
C|Superfamily: AMP-activated protein kinase; protein kinase homology
C|Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F:17-271/Domain: protein kinase homology <KIN>
F:25-33/Region: protein kinase ATP-binding motif
F:48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
F:147,151/Binding site: magnesium (Aen, Asp) #status predicted

Query Match          14.6%; Score 276; DB 1; Length 511;
Best Local Similarity 28.3%; Pred. NO.1.2e-11;
Matches 86; Conservative 49; Mismatches 115; Indels 54; Gaps 10;

Qy      80 GRAYQALHCPCTGTEYKCV-----PVQALAVLEPYARLPKHVAPTEVL 127
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Dy      31 GKVKIAEHLTLGKHAVKILNRKKIKMEEMEEVREIKIL-----RLFMHPIIRLYEVV 86

Qy      128 AGTQLLYAFETTR-NHDMSLVTRNRIPEPEAAVLFRQATALACHONGVLRLDKLC 186
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Dy      87 ETPSDVIYVMEYKSGLEPDYIVKEKGLQDEAKRKFQQLISGVEYCHRMVYNRRDKPE 146

Qy      187 RFVFADREKRLV---LEN-----LEDSCLVLPDSDIMDKACRAVYGPETLSSRAS 236
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Dy      147 NLLLDKRMVKNKINDFGLSINMRDGHFLKTSK-----GSPNVAAREVISGKL- 192

Qy      237 YSGKADWMSLVALTMTLAGHVPQDSEFVLLFGKIRCAVAPLPGLSAPACIVRCIL 296

```

Db 193 YAGEVWVSCGVTIYALCGTLPEDDENIPNLFKKIKGMIISPSHLSAGARDLIPML 252

Qy 297 RREAEELITATGILLHPLRODPMPLAPTRSHLEAQAQVVDGGLDEAREED--REV 354

Db 253 IVDPMKMTIPEIMNHMF-----QMLPRYLAVPPTD-MQAKKIDEDIIQEV 301

Qy 355 VLYG 358

Db 302 VKRG 305

RESULT 13

T02496
probable protein kinase [imported] - Arabidopsis thaliana
N/Alternate names: hypochemical protein T19C21.2
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C/Accession: T02496; F84805
R/Roundley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Erandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, August 1998
A/Description: Arabidopsis thaliana chromosome II BAC T19C21 genomic sequence.
A/Reference number: Z14676
A/Accession: T02496
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-431 <ROU>
A/Cross-references: EMBL:AC004683; NID:G3395421; PID:G3395423
A/Experimental source: cultivar Columbia
R/Lin, X.; Kaul, S.; Roundley, S.D.; Shee, T.P.; Benito, M.T.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Mofat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
Nause, D.; Nlema, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; NCID:20083487; PMID:10617197
A/Accession: F84805
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-431 <STO>
A/Cross-references: GB:AE002093; NID:G3786023; PID:AA067369.1; GSPDB:GN00139
C/Genetics:
A/Genes: T19C21.2 Atg38490
A/Map position: 2
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

Query Match 14.5%; Score 274; DB 2; Length 431;
Best Local Similarity 29.5%; Pred. No. 1.4e-11;
Matches 92; Conservative 45; Mismatches 141; Indels 34; Gaps 9;

Qy 54 TAPDRATAVATSRGLGPPVLEPBEGRAYOALHCPGTGE-YTCVYP-----V 101

Db 16 TDDDNKSALEFGKYDLGK--LLGSAFAKYQAEEDLQNGESVAIKVQKRLKDLTAHV 73

Qy 102 GEAALVLEVYARLPKHKHARPEVLA-GTOLLVAFTRTHGDMSLVTRHRIPEEA 160

Db 74 KEISVM---RLRHPHVILSEVALATKTKYFWEALAKGGELEPSR-TSNRFESLSR 128

Qy 161 VLEFQWATLAHQHGLVLRDLKCRFYADREKKVLLENEDSCVLGTGDDSLMDH 220

Db 129 KFEKQLISAVRYCHARGVHRDLKPEMLLDNRDLKVSDFLSAMKEQIHDPGMLHTTC 188

Qy 221 ACPAYVGEIILSSRASYSKADWLSGLVAFMTLAGHYFPDQSPVLLFGKIRGAVAL 280

Db 189 GTPAVVAEELLLK--GYDGSKADIMSCGVFLNAGYLPFRDPRIMGILYKTHIAQYKL 247

Qy 281 PAGSAPARCVRCILREPAERLTATGILLHPMLRODPMPLAPTRSHLEAQAQVVDG 340

Db 248 PDWTSDDRKLRLRLBPPELITVEETLKDPWF-----NHGVDPSEIT--GI 294

Qy 341 GLDEAREEEDR 352

Db 295 QADVDLLENCK 306

RESULT 14

A26030
serine/threonine-specific protein kinase (EC 2.7.1.1-) SNF1 - Yeast (Saccharomyces cerev.
N/Alternate names: protein YDR477w
C/Species: Saccharomyces cerevisiae
C/Date: 20-Aug-1987 #sequence_revision 20-Aug-1987 #text_change 23-Mar-2001
C/Accession: A26030; S69644
R/Celenza, J.U.; Carlson, M.
Science 233, 1175-1180, 1986
A/Title: A yeast gene that is essential for release from glucose repression encodes a p
A/Reference number: A26030; NCID:86289463; PMID:3526554
A/Accession: A26030
A/Molecule type: DNA
A/Residues: 1-633 <CEL>
A/Cross-references: EMBL:M13971; NID:G172629; PID:AAA35058.1; PID:G172630
R/Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A/Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.
A/Reference number: S69554
A/Accession: S69644
A/Molecule type: DNA
A/Residues: 1-633 <DIE>
A/Cross-references: EMBL:U33050; NID:G927726; PID:AA064904.1; PID:G927732; GSPDB:GN000C
C/Genetics:
A/Genes: SGD:SNF1; MIPS:YDR477w
A/Cross-references: SGD:S0002885; MIPS:YDR477w
A/Map position: 4R
C/Function:
A/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
A/Note: required for expression of glucose-repressed genes in response to glucose depriv
C/Superfamily: AMP-activated protein kinase; protein kinase homology
C/Keywords: ATP; autophosphorylation; magnesium; nucleus; phosphoprotein; phosphotransfe
F;33-306/Domain: protein kinase ATP-binding motif
F;51-69/Region: protein kinase ATP-binding motif
F;84,103,177,179/Active site: Lys, Glu, Asp, Lys #status predicted
F;182,186/Binding site: magnesium (Asn, Asp) #status predicted
F;210/Binding site: phosphate (Thr) (covalent) (by unidentified kinase) #status predictce

Query Match 14.5%; Score 273.5; DB 1; Length 633;
Best Local Similarity 27.8%; Pred. No. 2.3e-11;
Matches 88; Conservative 53; Mismatches 138; Indels 37; Gaps 11;

Qy 51 SEPTAPDRATAVATSRGLGPPVLEP-BEG--GRAYOALHCPGTGEYTCVYPVDGALV 107

Db 35 SSTLNPKSLADAGHIGNYQIVKTLGSGRGKYLAAHTTGQKVALKIIN-KVLLAK 93

Qy 108 LE-----PYARLPKHKHARPEVLAGTQLLYAFTRTHGDMSLVTRHRIPEE 158

Db 94 SPMQRIEREISYLLKLRHPIIKLYDVTKSDELTIMVEYAGNELPDYIVDRKMSBOE 153

Qy 159 AAVLFQWATLAHQHGLVLRDLKCRFYADREKKVLLENEDSCVLGTGDDSLMD 218

Db 154 ARPFQOQIISAVEYCHRHKIVARDLKPENILLDENLVKIDAFGL--SNIMT---DGNFL 208

Qy 219 KHAC--PAYVGEIILSSRASYSKADWLSGLVAFMTLAGHYFPDQSPVLLFGKIRG 276

Db 209 KTCGSEFYIADPEVLSGL-YAGEVVDWVSCGVILLYVLCRLPDDDSIPVLFNINS 267

Qy 277 AYALPAGSAPARCVRCILREPAERLTATGILLHPMLROD-PMPLAPTRSHLEAQAQ 335

Db 268 VYTLPEKFLSPGAAGIKMLVLPNLRISIHIMQDDMFKVDLPFYLLP----- 316

Qy 336 VPDGIGLDEAREEED 351

Db 317 -PD--LKPHEEENE 328

RESULT 15

S60304
serine/threonine-specific protein kinase (EC 2.7.1.1-) BKIN12 (version 2) - barley
C/Species: Hordeum vulgare (barley)
C/Date: 19-Mar-1997 #sequence_revision 15-Aug-1997 #text_change 11-Jun-1999

R;Accession: S60304; S24579
R;Halford, N.G.; Vicente-Carbajosa, J.; Sabelli, P.A.; Shewry, P.R.; Hammappel, U.; Kretzschmar, H. 1992
Plant J. 2, 791-797, 1992
A;Title: Molecular analyses of a barley multigene family homologous to the yeast protein tyrosine phosphatase Ptp1p
A;Reference number: S60303; MUID:93258420; PMID:1302632
A;Accession: S60304
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-513 <HAL>
A;Cross-references: EMBL:X65604
R;Halford, N.G.
submitted to the EMBL Data Library, April 1992
A;Reference number: S24578
A;Accession: S24579
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-61, 'A', 63-513 <HA>
A;Cross-references: EMBL:X65604; NID:G18933; PIDN:CAA6554.1; PID:G18934
C;Function:
A;Description: catalyzes the formation of peptidyl-L-serine-phosphate or peptidyl-L-threonine-phosphate
C;Superfamily: AMP-activated protein kinase; protein kinase homolog
C;Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F;15-272/Domain: protein kinase homolog <KIN>
F;23-31/Region: protein kinase ATP-binding motif
F;46,65,143,145/Active site: Lys, Glu, Asp, Lys #status predicted
F;148,157/Binding site: magnesium (Asn, Asp) #status predicted

Query Match	Score	DB 1	Length
14.4%	273	513	

Best Local Similarity 31.7%; Pred. No. 2e-11;
Matches 79; Conservative 42; Mismatches 102; Indels 26; Gaps 7

QY 117 HKIVARPREVLVAGTQLYAFETR-DHGDMHSLVTRRNRIRPEEAAVLFGOMATYALAHGQ 175
D6 77 HPHILIRVEVELETTPKQIFVMEYCNNGELDIYITIENGRLOEDBARIRIFQIILAGVGYCHR 136
QY 176 HGLVLRDYLKLRFPVADBERKKLV---LENLEDSCVLIGPDDSLMDKHAQPA--YVGPEI 230
D6 137 IMVYHRDILKPERNILLDSRYNVKLADFGLSNV-----MRDGHFLKTSQGSINLYAPAEI 188
QY 231 LSSRASVSGCKADVWSLGVALTFTMLAGHYFPQDSEPVULLFGKIRRAVYLLPAGLSAPARC 290
D6 189 ISSKK-YAGPEVDVWMSGVVLATLQGVSPFDDDNIPSLFRKIKGGTYILLPSTLSASAK 247
QY 291 LVRCILRREPAPRLTATGILHPWLROD--EMPLAPTRSHIMEAAQVVPDGLGDEAREEE 349
D6 248 LIPKLTINIDPMKRITIFHEIRVHPWFKNHLPCYLAVDPYPKCOAKMI-----DED 297
QY 350 GDREYVLYNG 358
D6 298 ILREYVNLG 306

Search completed: August 24, 2004, 18:53:44
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 18:49:56 ; Search time 12 Seconds
(without alignments)
1553.426 Million cell updates/sec

Title: US-10-070-337-5

Perfect score: 1892
Sequence: 1 MRATPLAPAGSLSRKKRLK.....GLGDFAREERGDEVVLYG 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1888	99.8	358	1 NIPK_HUMAN	Q96VU7 homo sapien
2	1338	70.7	354	1 NIPK_MOUSE	Q8K4K2 mus musculu
3	1324.5	70.0	349	1 NIPK_RAT	Q9W4G6 rattus norv
4	308	16.3	619	1 SNF1_CANTR	Q94168 candida ttr
5	305.5	16.1	646	1 CNK_HUMAN	Q9H4B4 homo sapien
6	303	16.0	620	1 SNF1_CANAL	P52497 candida alb
7	299.5	15.8	631	1 CNK_MOUSE	Q60906 mus musculu
8	294	15.5	794	1 KII1_HUMAN	Q9R011 rattus norv
9	292	15.4	615	1 CNK_RAT	Q8TDC3 homo sapien
10	284	15.0	1518	1 KKK1_YEAST	Q9R011 rattus norv
11	283.5	15.0	611	1 SNF1_CANGA	P34244 saccharomyc
12	280	14.8	535	1 KII1_ARATH	Q00372 candida gla
13	273.5	14.5	633	1 SNF1_YEAST	Q38997 arabidopsis
14	270.5	14.3	593	1 CDRI_SCHPO	P07334 schizosacch
15	270	14.3	576	1 SNF1_SCHPO	O74536 schizosacch
16	268.5	14.2	533	1 KCCD_RAT	P15791 rattus norv
17	268.5	14.2	736	1 KCCB_MOUSE	O81W33 homo sapien
18	266	14.1	542	1 KCCB_HUMAN	P28652 mus musculu
19	264	14.0	622	1 YNA3_CABEL	P45894 caenorhabdi
20	263	13.9	542	1 KCCB_RAT	P08413 rattus norv
21	263	13.9	542	1 KCCB_HUMAN	Q13554 homo sapien
22	262	13.8	1142	1 GIN4_YEAST	Q12263 saccharomyc
23	261.5	13.8	499	1 KCCD_HUMAN	Q13557 homo sapien
24	260.5	13.8	552	1 AAK2_HUMAN	P54646 homo sapien
25	259.5	13.7	552	1 AAK2_RAT	O09137 rattus norv
26	259.5	13.7	891	1 KINI_SCHPO	P22987 echinosacch
27	257.5	13.6	786	1 SNIL_HUMAN	P57059 homo sapien
28	257	13.6	502	1 KRI1_SECCB	Q02703 secale cere
29	255	13.5	470	1 CDB1_ARATH	Q06850 arabidopsis
30	251.5	13.3	478	1 KCCA_HUMAN	O9UGM7 homo sapien
31	251.5	13.2	433	1 KCCA_RAT	P11275 rattus norv
32	249.5	13.1	478	1 STXB_HUMAN	Q15831 homo sapien
33	248.5	13.1	685	1 SNK_HUMAN	Q9NYJ3 homo sapien

34	246.5	13.0	353	1 ASK2_ARATH	P43292 arabidopsis
35	246	13.0	332	1 AAI2_WHEAT	Q02066 triticum ae
36	245	12.9	735	1 K6A1_RAT	O63531 rattus norv
37	243.5	12.9	472	1 KCCG_HUMAN	O13555 homo sapien
38	243.5	12.9	527	1 KCCG_RAT	P11730 rattus norv
39	243.5	12.9	39	1 KCCG_MOUSE	Q92369 mus musculu
40	241.5	12.8	724	1 K6A1_MOUSE	P18653 mus musculu
41	241	12.7	682	1 SNK_MOUSE	P53351 mus musculu
42	241	12.7	1037	1 KCCA_YEAST	P25189 saccharomyc
43	240.5	12.7	478	1 KCCA_MOUSE	P11798 mus musculu
44	240.5	12.7	548	1 AAK1_RAT	P54645 rattus norv
45	240	12.7	542	1 CDP3_ORYSA	P53684 oryza sativ

ALIGNMENTS

RESULT 1
NIPK_HUMAN STANDARD. PRT. 358 AA.
ID NIPK_HUMAN Q96RUT; Q9H5M8; Q9HND2;
AC Q96RUT; Q9H5M8; Q9HND2;
DT 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neuronal cell death inducible putative kinase (SKIP3).
GN NIPK OR C20ORP97.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RA Kiss-Toth E., Wyllie D.H., Qvarnstrom E.E., Dower S.K.;
RT Identification of pro-inflammatory cytokine signalling network
RT components by transcription expression screening."
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Oca T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NPO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Garder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cordy N.R.,
RA Ellington A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Eilington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Gatliff D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.U., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leveson-Jones M.H., Leverhame M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McKay K., McMurtry A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver R.K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Pratchalingam S.R., Plumb R.W., Ramay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showmken R., Sims S.,
RA Shuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RL "The DNA sequence and comparative analysis of human chromosome 20.";
Nature 414:665-671(2001).

[4]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Cervix, and Muscle;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.U., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: May play an important role in a common pathway leading
 CC to programmed neuronal cell death. Does not appear to function in
 CC the programmed death of non-neuronal cells. May serve as an
 CC endogenous antagonist competing for substrate with functional
 CC kinases that act to promote neuronal cell survival (By
 CC similarity).
 CC -1- SIMILARITY: Contains 1 protein kinase domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF250311; AAK58175.1; -
 CC EMBL; AK026945; BAB15597.1; -
 CC EMBL; AL034548; CAB81634.1; -
 CC EMBL; BC019363; AAH19363.1; -
 CC EMBL; BC027484; AAH27484.1; -
 CC Genew; HGNC:16228; C20orf97.
 CC MIM; 607898; -
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR002290; Ser_thr_kinase.
 CC Pfam; PF00069; pkinase; 2.
 CC ProDom; PD000001; Prot_kinase; 1.
 CC SMART; SM00220; S_TKc; 1.
 CC DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 CC DR KIM Apoptosis.
 CC KIM DOMAIN 68 316 PROTEIN KINASE.
 CC FT CONFLICT 84 84 Q -> R (IN REF. 4; AAH27484).
 CC FT CONFLICT 105 105 L -> P (IN REF. 1).
 CC FT CONFLICT 114 114 L -> V (IN REF. 2).
 CC FT CONFLICT 194 195 ER -> DREK (IN REF. 1).
 CC SEQUENCE 358 AA; 39577 MW; CEISFBD9A81BD63 CRC64;
 SQ
 Query Match 99.8%; Score 1888; DB 1; Length 358;
 Best Local Similarity 99.7%; Pred. No. 7.5e-139;
 Matches 357; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 121 ARPTEVAGTOLLYAFTRTHGDMHSIVTRHRIPEEPAAVLFROMATATAHCHQHGVL 180
 DB 121 ARPTEVAGTOLLYAFTRTHGDMHSIVSRHRIPEEPAAVLFROMATATAHCHQHGVL 180
 QY 181 RDLKLCFVFPADRRKKLVLENLEDSCVLGGPDSDMDKACPAVYGEITSSASISGK 240
 DB 181 RDLKLCFVFPADRRKKLVLENLEDSCVLGGPDSDMDKACPAVYGEITSSASISGK 240
 QY 241 AADVMSIGVALFTMLAGHYFPODESEVPLTFGKIRGAYALPAGLSAPARCTVCLLRREP 300
 DB 241 AADVMSIGVALFTMLAGHYFPODESEVPLTFGKIRGAYALPAGLSAPARCTVCLLRREP 300
 QY 301 AERLTATGILLHFWLRDDPEPLAPTRSHLWEAAQVDPDGLDDEAREEGDREVLVYG 358
 DB 301 AERLTATGILLHFWLRDDPEPLAPTRSHLWEAAQVDPDGLDDEAREEGDREVLVYG 358
 RESULT 2
 NIPK_MOUSE STANDARD; PRT; 354 AA.
 AC Q8K4K2; Q921E7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neuronal cell death inducible putative kinase (TRB-3).
 GN NIPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kiss-Toth E., Dempsey C., Jozsa V., Caunt J., Oxley K.M.,
 RA Bagstaff S.M., Wyllie D.H., Harre M., O'Neill L.A.J., Qvarnstrom E.E.,
 RA Dower S.K.;
 RT "Mammalian homologs of Drosophila tribbles (trtb) control mitogen
 RT activated protein kinase signaling".
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Otsu N., Saito R., Suzuki H., Yamataka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotojori T.,
 RA Balazselli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chochla C., Corbani L.E., Cousins S.,
 RA Dalia E., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Petre G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.O., Reid J., Ring B.Z., Ringwald M.,
 RA Sanderlin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yamagisawa M., Yang J., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Komio H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Maki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs".
 RL Nature 420:563-573 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Maruina K., Farmer A.F., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halysk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.T., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -i- FUNCTION: May play an important role in a common pathway leading
 CC to programmed neuronal cell death. Does not appear to function in
 CC the programmed death of non-neuronal cells. May serve as an
 CC endogenous antagonist competing for substrate with functional
 CC kinases that act to promote neuronal cell survival (by
 CC similarity).
 CC -i- SIMILARITY: Contains 1 protein kinase domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF588668; AAM45476.1; -;
 CC EMBL: AK089931; BAC41002.1; -;
 CC EMBL: BC012955; AAH12955.1; -;
 CC InterPro: IPR000719; Prot_kinase.
 CC InterPro: IPR002290; Ser_thr_kinase.
 CC Pfam: PF00069; pkinase; 2.
 CC InterPro: PD000001; Prot_kinase; 1.
 CC SMART: SM00220; S_TKc; 1.
 CC PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
 CC KW Apoptosis.
 CC FT DOMAIN 68 315 PROTEIN KINASE.
 CC FT CONFLICT 157 219 S -> P (IN REF. 3).
 CC FT CONFLICT 219 219 K -> T (IN REF. 1).
 CC FT CONFLICT 239 264 MISSING (IN REF. 3).
 CC FT CONFLICT 301 354 SERVALGILHPMLREDHGKRVSPQSDREMDQVVDGPGQ
 CC LEDEAGEGVGNG -> CRATCGPGNPLMSIVRGRSPSL
 CC STY (IN REF. 2).
 CC SEQUENCE 354 AA; 39022 MW; 2CB283FC19F859F CRC64;
 SQ
 Query Match 70.7%; Score 1338; DB 1; Length 354;
 Best Local Similarity 73.5%; Pred. No. 2.5e-96;
 Matches 263; Conservative 31; Mismatches 60; Indels 4; Gaps 2;
 QY 1 MRATPLAAPAGSISRKKRLLELDNLTDPVQKRRASGQFRLPCLPLSPPTADPRAT 60
 DB 1 MRATPLAAPAGSISRKKRLLELDNLTDPVQKRRASGQFRLPCLPLSPPTADPRAT 60
 QY 61 AVATASRLGTYVLLLEPEGGRAYOALHCPGTGYTCVKVYVQDALVLEPYALLPKHV 120
 DB 61 AVATASRLGTYVLLLEPEGGRAYOALHCPGTGYTCVKVYVQDALVLEPYALLPKHV 120
 QY 121 ARPTVLAGTQLLYAFRTTHGDMHSLVTRRHIPPEFAAVLFRQMATALAHCHQGLVL 180
 DB 121 ARPTVLAGTQLLYAFRTTHGDMHSLVTRRHIPPEFAAVLFRQMATALAHCHQGLVL 180
 QY 121 ARPTVLAGTQLLYAFRTTHGDMHSLVTRRHIPPEFAAVLFRQMATALAHCHQGLVL 180
 DB 121 ARPTVLAGTQLLYAFRTTHGDMHSLVTRRHIPPEFAAVLFRQMATALAHCHQGLVL 180
 QY 181 RDLKLCRFVPADEKRRKLVLENLEDSCVLTGPPDSDLMDKACPAVYVGPILLSSRPSYSGK 240
 DB 181 RDLKLCRFVPADEKRRKLVLENLEDSCVLTGPPDSDLMDKACPAVYVGPILLSSRPSYSGK 240

DB 181 RDLKLCRFVPADEKRRKLVLENLEDSCVLTGPPDSDLMDKACPAVYVGPILLSSRPSYSGK 240
 QY 241 AADVMSIGVALFTMLAGHYPPQDSEPVLLRGKTRRGAYALPAGISAPARCLVRCILRREP 300
 DB 241 AADVMSIGVALFTMLAGHYPPQDSEPVLLRGKTRRGAYALPAGISAPARCLVRCILRREP 300
 QY 301 AERLTATGILHPMLRQDPMPLAETRSHLWMAQVVDGGLDARREEGREVLVXG 358
 DB 301 SERVALGILHPMLREDHGKRVSPQSDREMDQVVDGPGQLEER-EEG--EVGLVIG 354
 RESULT 3
 NIPK_RAT
 ID NIPK_RAT STANDARD; PRT; 349 AA.
 AC 99706;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neuronal cell death inducible putative kinase.
 GN NIPK.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Neuron;
 RX MEDLINE=99262087; PubMed=10329375;
 RA Mayumi-Macuda K., Kojima S., Suzuki H., Sakata T.;
 RT "Identification of a novel kinase-like gene induced during neuronal
 RT cell death.";
 RL Biochem. Biophys. Res. Commun. 258:260-264(1999).
 CC -i- FUNCTION: May play an important role in a common pathway leading
 CC to programmed neuronal cell death. Does not appear to function
 CC in the programmed death of non-neuronal cells. May serve as an
 CC endogenous antagonist competing for substrate with functional
 CC kinases that act to promote neuronal cell survival.
 CC -i- TISSUE SPECIFICITY: Detected only in the lung. Not detected in the
 CC heart, brain, spleen, liver, skeletal muscle, kidney and testis.
 CC -i- INDUCTION: Expression induced during programmed cell death evoked
 CC in neuronal cells by NGF-depletion.
 CC -i- SIMILARITY: Contains 1 protein kinase domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB020967; BAA77582.1; -;
 CC InterPro: IPR000719; Prot_kinase.
 CC Pfam: PF00069; pkinase; 2.
 CC ProDom: PD000001; Prot_kinase; 1.
 CC PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
 CC KW Apoptosis.
 CC FT DOMAIN 63 310 PROTEIN KINASE.
 CC FT CONFLICT 157 219 S -> P (IN REF. 3).
 CC FT CONFLICT 219 219 K -> T (IN REF. 1).
 CC FT CONFLICT 239 264 MISSING (IN REF. 3).
 CC FT CONFLICT 301 354 SERVALGILHPMLREDHGKRVSPQSDREMDQVVDGPGQ
 CC LEDEAGEGVGNG -> CRATCGPGNPLMSIVRGRSPSL
 CC STY (IN REF. 2).
 CC SEQUENCE 349 AA; 38602 MW; 3050F9BF8346D815 CRC64;
 SQ
 Query Match 70.0%; Score 1324.5; DB 1; Length 349;
 Best Local Similarity 72.9%; Pred. No. 2.7e-95;
 Matches 261; Conservative 28; Mismatches 60; Indels 9; Gaps 3;
 QY 1 MRATPLAAPAGSISRKKRLLELDNLTDPVQKRRASGQFRLPCLPLSPPTADPRAT 60
 DB 1 MRATPLAAPAGSISRKKRLLELDNLTDPVQKRRASGQFRLPCLPLSPPTADPRAT 60
 QY 61 AVATASRLGTYVLLLEPEGGRAYOALHCPGTGYTCVKVYVQDALVLEPYALLPKHV 120
 DB 61 AVATASRLGTYVLLLEPEGGRAYOALHCPGTGYTCVKVYVQDALVLEPYALLPKHV 120
 QY 121 ARPTVLAGTQLLYAFRTTHGDMHSLVTRRHIPPEFAAVLFRQMATALAHCHQGLVL 180
 DB 121 ARPTVLAGTQLLYAFRTTHGDMHSLVTRRHIPPEFAAVLFRQMATALAHCHQGLVL 180
 QY 121 ARPTVLAGTQLLYAFRTTHGDMHSLVTRRHIPPEFAAVLFRQMATALAHCHQGLVL 180
 DB 121 ARPTVLAGTQLLYAFRTTHGDMHSLVTRRHIPPEFAAVLFRQMATALAHCHQGLVL 180

```

Db 116 ARPTEVILGSQLTYFTFKTHGDLHSIVRSRGIPEPEAAALFRQMSAVAHCHKHGL 175
QY 181 RDLKCFEVPFADRRKKLVLENLEDSCLVTPGPDSDMDKACPAVYGPETISSASYSCK 240
Db 176 RDLKRFVSNCRKTLVLENLEDSCLVTPGPDSDMDKACPAVYGPETISSASYSCK 235
QY 241 AADVMSIGVALFTMLAGHYPPQDSEPVLLFGKIRGAYVALPAGISAPARCIVRCILRRP 300
Db 236 AADVMSIGVALFTMLAGHYPPQDSEPVLLFGKIRGAYVALPAGISAPARCIVRCILRRP 295
QY 301 AERTLTAGTILHPELRQDPMPLATRSHLMFAAQQVVDGILGDFAREEGDREVLYNG 358
Db 296 SERLVALGILHPELRQDPMPLATRSHLMFAAQQVVDGILGDFAREEGDREVLYNG 349

```

RESULT 4

```

SNP1_CANTR STANDARD; PRT; 619 AA.
ID 094168;
AC 094168;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
GN SNP1.
OS Candida tropicalis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
NCBI_TaxID=5482;
RN [1]

```

```

RA Kanal T., Ogawa K., Ueda M., Tanaka A.;
RT "Genetic evaluation of the function of SNP1 in Candida tropicalis.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Essential for release from glucose repression. It
CC interacts and has functional relationships to the regulatory
CC protein SNF4. Could phosphorylates CAT8 (By similarity).
CC -1- SUBCELLULAR LOCATION: Associated with the nuclear membrane (By
CC similarity).
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNP1
CC subfamily.

```

```

CC This SNF1-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; AB024535; BAA75889.1; -.
DR HSSP; C63450; 1A06.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser_Thr_kin_AS.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE_SF; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Phosphorylation; Serine/threonine-protein kinase; ATP-binding;
KW Transphosphorylation; Carbohydrate metabolism; Nuclear protein.
FT DOMAIN 17 29
FT DOMAIN 303 303
FT NP_BIND 58 66
FT BINDING 81 81
FT ACT_SITE 174 174
FT ACT_SITE 207 207
FT MOD_RES 207
SQ SEQUENCE 619 AA; 70323 MW; 0FCF1FC3DCE706D7 CRC64;

```

```

Query Match 16.3%; Score 308; DB 1; Length 619;
Best Local Similarity 30.9%; Pred. No. 2.1e-16;

```

```

Matches 95; Conservative 48; Mismatches 128; Indels 36; Gaps 12;
QY 40 QPRLPCLPLSPPTADPRATAVATASRLGPEYLLER-EEG--GRAYQALHCPGTETC 96
Db 32 QPAPQ--IPIDPNVND-----ANRIGRYQIILKTIGESSFKVYLAQHVGVGQVAL 80
QY 97 KVVYQVQALAVLE-----PYARLPKPKHVARPFEVYLAGTOLLYAFRTKGDHSL 147
Db 81 KIIN-RKTLKSDQGVVEREISTRLRHPHIIKTDVYIKSKDEIIMVIEFAPKKEFDY 139
QY 148 VRTRRRIPEPEAAVLFROMATALAHCHQGLVLRDLKCFEVPFADRRKKLVLENLEDSCL 207
Db 140 IVRGKMPDEARFRFPQIIAAVEYCHRHKIYRDLKPEMLDDQLANVKIADFGI--SN 197
QY 208 VLTGPDSDMDKAC--PAYVGPETISSASYSCKADVMSIGVALFTMLAGHYPPQDSE 265
Db 198 IMT---DGNFLKTSKSGSPNYAAPEVISGKL-YAGPEVDVMSGVYLYVMCGRLPFDEF 253
QY 266 PVLLFGKIRGAYVALPAGISAPARCIVRCILRRPAPRLTATGILLHPELRQDPM--L 322
Db 254 IPELFKXISGVYTLDPNVYSPGAKHLLTRMLVVPNPLRITIHIEDEWFKOD-MPDYIL 312
QY 323 APTRSHL 329
Db 313 PPDLSKI 319

```

RESULT 5

```

CNK_HUMAN STANDARD; PRT; 646 AA.
ID 09H434; O15767;
AC 09H434; O15767;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cytokine-inducible serine/threonine-protein kinase (EC 2.7.1.37) (FGF-
DE inducible kinase) (Proliferation-related kinase).
GN CNK OR FNK OR PRK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Embryo.
RX MEDLINE=20493044; PubMed=11039900;
RA Holtrich U., Wolf G., Yuan J., Bereiter-Hahn J., Karn T., Weiler M.,
RA Kauselmann G., Rehli M., Andreesen R., Kaufmann M., Kuhl D.,
RA Strebhardt K.;
RT "Adhesion induced expression of the serine/threonine kinase Fnk in
RT human macrophages.";
RL Oncogene 19:4832-4839(2000).
RN [2]
RP SEQUENCE OF 28-646 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=96325053; PubMed=8702627;
RA Li B., Ouyang B., Pan H., Reissmann P.T., Slamon D.J., Arceci R.,
RA Lu L., Dai W.;
RT "Prk, a cytokine-inducible human protein serine/threonine kinase whose
RT expression appears to be down-regulated in lung carcinomas.";
RL J. Biol. Chem. 271:19402-19408(1996).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=98019242; PubMed=9353331;
RA Ouyang B., Pan H., Lu L., Li U., Stambrook P., Li B., Dai W.;
RT "Human Prk is a conserved protein serine/threonine kinase involved in
RT regulating M phase functions.";
RL J. Biol. Chem. 272:28646-28651(1997).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=20027391; PubMed=10557092;
RA Ouyang B., Li W., Pan H., Meadows J., Hoffmann I., Dai W.;
RT "The physical association and phosphorylation of Cdc25C protein
RT phosphatase by Prk.";

```

RL Onco gene 18:6029-6036(1999).
 CC -1- FUNCTION: Serine/threonine protein kinase involved in regulating M
 CC phase functions during the cell cycle. May also be part of the
 CC signaling network controlling cellular adhesion. In vitro, is able
 CC to phosphorylate CDC25C and casein.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SUBUNIT: Binds to the calcium/integrin-binding protein (CIB). This
 CC interaction probably occurs via the POLO-box domain.
 CC -1- SUBCELLULAR LOCATION: Membrane-associated.
 CC -1- TISSUE SPECIFICITY: Transcripts are highly detected in placenta,
 CC lung, followed by skeletal muscle, heart, pancreas, ovaries and
 CC kidney and weakly detected in liver and brain. May have a short
 CC half-life, in cells of hematopoietic origin, strongly and
 CC exclusively detected in terminally differentiated macrophages.
 CC Transcript expression appears to be down-regulated in primary
 CC lung tumor.
 CC -1- INDUCTION: Cytokine and cellular adhesion trigger FAK induction.
 CC -1- PM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS
 CC CELLS EXIT MITOSIS (BY SIMILARITY).
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 CC CDC5/Polo subfamily.
 CC -1- SIMILARITY: Contains 2 POLO box domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ293866; CAC10659.1; -.
 CC DR EMBL: U56998; AAC50637.1; ALT_INIT.
 CC DR Genew; HGNC:2154; CNK.
 CC DR GK; Q94B4; -.
 CC DR MIM; 602913; -.
 CC DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
 CC DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
 CC DR GO; GO:0000074; P:regulation of cell cycle; TAS.
 CC DR InterPro; IPR000959; POLO box.
 CC DR InterPro; IPR000719; Prot_kinase.
 CC DR InterPro; IPR008271; Ser_thr_pkin_AS.
 CC DR InterPro; IPR002290; Ser_thr_pkinase.
 CC DR Pfam; PF00069; pkinase; 1.
 CC DR Pfam; PF00659; POLO box; 2.
 CC DR ProDom; PD000001; Prot_kinase; 1.
 CC DR SMART; SM00220; S_TKc; 1.
 CC DR PROSITE; PS50078; POLO_BOX; 2.
 CC DR PROSITE; PS50107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 CC KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
 CC phosphorylation.
 CC FT DOMAIN 62 314 PROTEIN KINASE.
 CC FT NP BIND 68 76 ATP (BY SIMILARITY).
 CC FT BINDING 91 91 ATP (BY SIMILARITY).
 CC FT ACT SITE 185 185 BY SIMILARITY.
 CC FT DOMAIN 470 537 POLO BOX 1.
 CC FT DOMAIN 567 637 POLO BOX 2.
 CC FT CONFLICT 99 99 V -> A (IN REF. 2).
 CC FT CONFLICT 353 353 V -> G (IN REF. 2).
 CC FT CONFLICT 419 419 H -> D (IN REF. 2).
 CC FT CONFLICT 464 470 FSEWGVF -> VSKWVDY (IN REF. 2).
 CC FT CONFLICT 522 522 R -> P (IN REF. 2).
 CC FT SEQUENCE 646 AA; 71789 MW; C20147CD0F8A3B4 CRC64;
 QY Query Match 16.1%; Score 305.5; DB 1; Length 646;
 QY Best Local Similarity 28.9%; Pred. No. 3.5e-16;
 QY Matches 101; Conservative 46; Mismatches 134; Indels 67; Gaps 15;
 DB 11 RPFQRTAAATAPAGPBP--PFSALRGPELEMLAGLFTSDPGRLLTPRSGRTYLLKGRLL 68

QY 69 GPVLLPEBERG--RAYQALHCPRTGYTCKVYVQDALVLEPYARLP----- 115
 DB 69 G-----KGGFRCREADTDTETGSAVAVKVPQSR---VVKHOREKINIELHRLD 117
 QY 116 PHKVARPTVLAAGTOLVYAFETR-THGDMHSLVTRHRIPEPAALVFROMATALACH 174
 DB 118 QHRITVPSHHFFEDADNIVITFLEICSKSLAHMKAKHTLLEPRVRYLRQILSGIKYLH 177
 QY 175 QHGLVLDLKLRFVPADRERKKVLNLEDSCVLTGPDSDLMDKAC--PAYVGPETLS 232
 DB 178 QRLTHHDLKLGNFET--ENMELKVGDFGLAARLEPEQR--KKTICGTNNVAAFEVLL 233
 QY 233 SRASYSGKADVWSLGVALLFTMLAGHPQDSEVLLFGKRRGAYALPAGISAPARCLV 292
 DB 234 RQG--HGPEADVWSLGCVMYTLTGSPPEFADUKETRYCRICKQVHYTLPASLSPARQL 291
 QY 293 RCLRRPFAERLVTATGTLHPMLRQ---DPMPIAPTRSHLMEAAQVVPD 338
 DB 292 AAILRASPDRPSIDQLRHDFFTKGYTPDKPLIS-----SCVTVVPD 333
 RESULT 6
 SNF1_CANAL STRAND: PRT; 620 AA.
 ID SNF1_CANAL
 AC P52497; Q00309;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
 GN SNF1.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 3254 / B-311;
 RX MEDLINE=98053924; PubMed=9393775;
 RA Pether R., Chang Y.C., Kwon-Chung K.U.;
 RT "A gene homologous to *Saccharomyces cerevisiae* SNF1 appears to be
 RT essential for the viability of *Candida albicans*.";
 RL Infect. Immun. 65:4909-4917(1997).
 RN [2]
 RP SEQUENCE OF 7-620 FROM N.A.
 RC STRAIN=ATCC 3254 / B-311;
 RA Pether R., Kwon-Chung K.U.;
 RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Essential for release from glucose repression. It
 CC interacts with and has functional relationship to the regulatory
 CC protein SNF4. Could phosphorylates Cat8 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Associated with the nuclear membrane (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L78129; AAB48643.1; -.
 CC DR EMBL: L39263; AAB92456.1; -.
 CC DR HSRP; Q63450; 1A06.
 CC DR InterPro; IPR000719; Prot_kinase.
 CC DR InterPro; IPR008271; Ser_thr_pkin_AS.
 CC DR InterPro; IPR002290; Ser_thr_pkinase.
 CC DR Pfam; PF00069; pkinase; 1.
 CC DR ProDom; PD000001; Prot_kinase; 1.
 CC DR SMART; SM00220; S_TKc; 1.

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
 KM Phosphorylation; Carbohydrate metabolism; Nuclear protein.
 FT DOMAIN 16 29
 FT NP BIND 53 305
 FT BINDING 59 67
 FT ACT SITE 175 175
 FT MOD RES 208 208
 FT CONFLICT 228 234
 FT CONFLICT 242 255
 FT CONFLICT 387 387
 FT CONFLICT 416 416
 FT CONFLICT 494 494
 SQ SEQUENCE 620 AA; 70005 MW; 1806c52b5061d2b CRC64;

Query Match
 Best Local Similarity 16.0%; Score 303; DB 1; Length 620;
 Matches 95; Conservative 55; Mismatches 131; Indels 42; Gaps 12;

QY 48 LPLSPPTAPDRAVAATASRLGPEVLLP-EEG--GRAVALHCPGTETCKVYPOEA 104
 DB VPIDPAMP-----ANRIGRYQLIKLTGEGSFGKVLQHLGQKVALKIN-RTK 88
 QY 105 LAYLE-----PYARLPHKHVARPTEVLAAGTQLLVAFRTIGDMHSLVTRHRIP 155
 DB LASSDMQGRVERISTURLRHRPHITKLVDVSKDEIIVWIEFAGEKEIDYIVQSGKMP 148
 QY 156 EPEAAVLFRMATATLAHQHGLVLRDLKCRFVFADEREKRIVLNLEDSCVLTGEPDS 215
 DB EDARERFQOQIIAAYEYCHRHKIVHDKRENLLDDQANVKIADGGL--SNMT---DG 203
 QY 216 LMDKHC--PAYV-GEIILSRASISGKADWLSGLVFTMLAGHYPFQDSEPVLLFGK 272
 DB 204 NPLKTSGSPNVPAPAEVVISGKL-VAGPEYDVMSAGVILYVMLCGRLPDEDEIPALFKK 262
 QY 273 IRRGAVALPAGLSAPARCIVRCGLRREPRLATGILLHPMLKODPMLAPRSLHMA 332
 DB 263 ISNGVTLTPNYSAGAKHLITRMLVYNPLNRIITHEIMEDWFKOD-MP----- 310
 QY 333 AQVVPDGLGLDEAREEGDREVY 355
 DB 311 DYLPLPDLSKNKNSKIDVEDVI 333

RESULT 7
 CNK_MOUSE
 ID CNK_MOUSE STANDARD; PRT; 631 AA.
 AC Q60806; Q60822; Q9R009;
 DT 16-OCT-2001 (Rel. 40. Created)
 DT 16-OCT-2001 (Rel. 40. Last sequence update)
 DT 28-FEB-2003 (Rel. 41. Last annotation update)
 DE Cytokine-inducible serine/threonine-protein kinase (EC 2.7.1.37) (FGF-inducible kinase).
 GN CNK OR FNK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=NIH Swiss;
 RX MEDLINE=95247749; PubMed=7730342;
 RA Donohue P.V.; Alberts G.F.; Guo Y.; Winkles J.A.;
 RT "Identification by targeted differential display of an immediate early
 gene encoding a putative serine/threonine kinase.";
 RL J. Biol. Chem. 270:10351-10357(1995).
 RN [2]
 SEQUENCE OF 333-437 FROM N.A. (ISOFORM 2).

RC STRAIN=NIH Swiss;
 RA Kauselmann G.; Weiler M.; Kuhl D.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP FUNCTION, AND PHOSPHORYLATION.
 RX MEDLINE=98343954; PubMed=9677325;
 RA Chase D.; Peng Y.; Hershaw B.; Winkles J.A.; Longo D.L.; Ferris D.K.;
 RT "Expression and phosphorylation of fibroblast-growth-factor-inducible
 kinase (Fhk) during cell-cycle progression.";
 RL Biochem. J. 333:655-660(1998).
 CC -1- FUNCTION: Serine/threonine protein kinase involved in regulating M
 phase functions during the cell cycle. May also be part of the
 signaling network controlling cellular adhesion. In vitro, is able
 to phosphorylate CDC35C and casein (By similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SUBUNIT: Binds to the calcium/integrin-binding protein (CIB). This
 interaction probably occurs via the POLO-box domain (By
 similarity).
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q60806-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q60806-2; Sequence=VSP_004927;
 CC -1- TISSUE SPECIFICITY: Expressed in skin.
 CC -1- PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS
 CC CELLS EXIT MITOSIS.
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 CC -1- SIMILARITY: Contains 2 POLO box domains.
 CC -1- SIMILARITY: Contains 2 POLO box domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U21392; AAC52191.1; -
 DR EMBL; U22434; AAC52192.1; -
 DR EMBL; AF136586; AAF08369.1; -
 DR PIR; A57286; A57286.
 DR MGI; MGI:109604; Cnk.
 DR InterPro; IPR000959; POLO_box.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_kin AS.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00659; pkinase; 1
 DR Pfam; PF00659; POLO_box; 2.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS50078; POLO_BOX; 2.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 KW PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KM Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;
 KW Phosphorylation; Alternative splicing.
 FT DOMAIN 63 315
 FT NP BIND 69 77
 FT BINDING 92 92
 FT ACT SITE 186 186
 FT DOMAIN 455 518
 FT DOMAIN 552 622
 FT VARSPLIC 373 373
 FT CONFLICT 386 386
 SQ SEQUENCE 631 AA; 70012 MW; 20857341870BD1D2 CRC64;
 Query Match
 Best Local Similarity 15.8%; Score 299.5; DB 1; Length 631;
 Matches 98; Conservative 48; Mismatches 133; Indels 67; Gaps 14;

Db 71 ITGQKVAIKIVNEKLSSEVLKVEREIALI-----KLIHPHVILKLDVYENKKYLYVL 126
 QY 138 TR-THGDMHSLVTRRHRIPEPEAAVLFROMATALAHCHQGLVLRDILKLFVADSEK 136
 Db 127 EHVSGELFPIYLKKGRLTKREARKFRQIVSALDFCHSYSTICHRDLKPNLL-DEKN 184
 QY 197 KLVLENEDSCVLTPGDDSLMDKAC--PAYVGEIILSPRASYSGKAADVWSLGVLFM 254
 Db 185 NRIADFGMAVLQVG--DSILET--SCGSPHYACPEVIKGE-KYDGRADWMSGCVILFAL 240
 QY 255 LAGHYPRQDSEPVLLFKIRGAYALPAGISAPRCVRCILRREPARELTATGILLHPW 314
 Db 241 LVGLHLPDDNNLRQLLEKVRGVMHMFIPDCQSLRGMIEVEPEKRLSLBQIKHPW 300
 QY 315 L--RQDPMPV--APTRSHLWEAAQVVPDGLGD-----EARE 347
 Db 301 YLGGKHEDPCLEPAPER--VAMRSLPSNGELDPVLESMASLGCFRPRELHRELRS 357
 QY 348 EEGDREVLVY 357
 Db 358 EENOEKMTY 367

RESULT 9

CNK RAT STANDARD; PRT; 615 AA.

AC Q9R011; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytochrome-inducible serine/threonine-protein kinase (EC 2.7.1.37) (FGF-inducible kinase) (Fragment).
 GN CNK OR FNK.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99452760; PubMed=10523297;
 RA Kauselmann G., Weller M., Wulff P., Jessberger S., Konietzko U., Scafield J., Stambli U., Beretner-Hahn J., Streibhardt K., Kuhl D.;
 RT "The polo-like protein kinases Fnk and Snk associate with a Ca(2+)-and integrin-binding protein and are regulated dynamically with synaptic plasticity.";
 RL EMBO J. 18:5528-5539 (1999).
 CC -!- FUNCTION: Serine/threonine protein kinase involved in regulating M phase functions during the cell cycle. May also be part of the signaling network controlling cellular adhesion. In vitro, is able to phosphorylate CDC25C and casein (By similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- SUBUNIT: Binds to the calcium/integrin-binding protein (CIB). This interaction probably occurs via the POLO-box domain.
 CC -!- SUBCELLULAR LOCATION: WHEN INDUCED, IT TRANSLOCATES INTO THE DENDRITES OF ACTIVATED NEURONS.
 CC -!- TISSUE SPECIFICITY: Constitutively expressed in post-mitotic neurons.
 CC -!- INDUCTION: By the intense activity associated with seizures.
 CC -!- PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS CELLS EXIT MITOSIS (By similarity).
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 CC -!- SIMILARITY: Contains 2 POLO box domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC BMBL; AFI36584; AAF08367.1; --

DR InterPro: IPR000959; POLO box.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR008271; Ser_thr_pkin_AS.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00659; pkinase; 1.
 DR Pfam: PF00659; POLO box; 2.
 DR Prodom: PD000001; Prot_kinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PSS0078; POLO BOX; 2.
 DR PROSITE: PSS0107; PROTEIN KINASE ATP; 1.
 DR PROSITE: PSS0011; PROTEIN KINASE DOM; 1.
 DR PROSITE: PSS0108; PROTEIN KINASE ST; 1.
 KM Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat; Phosphorylation.
 FT NON TER 1
 FT DOMAIN 31 283 PROTEIN KINASE.
 FT NE BIND 37 45 ATP (BY SIMILARITY).
 FT BINDING 60 60 ATP (BY SIMILARITY).
 FT ACT SITE 154 154 BY SIMILARITY.
 FT DOMAIN 440 503 POLO BOX 1.
 FT DOMAIN 537 607 POLO BOX 2.
 FT NON TER 615
 SQ SEQUENCE 615 AA; 68800 MW; 06584C229B5D71BD CRC64;

Query Match 15.4%; Score 292; DB 1; Length 615;
 Best Local Similarity 28.5%; Pred. No. 3.6e-15;
 Matches 94; Conservative 47; Mismatches 127; Indels 62; Gaps 13;

QY 41 PRLPCLILSPETADRAATAVATASRLGPVLLPEEG-----GRAYQL 86
 Db 3 PSESEPV--LAGQTPD-----ASRL-----ITDRSGRTYIKGRLLGKGFPARYEN 49
 QY 87 HCPPTGETCYKYYPVGEALAVLEPYARLP-----PKHVAPPEVLAAGTOLLA 135
 Db 50 DRETSAVAVKXVLPQSR--VAKPHQREKIINEIELHRLDLOHRHIVRFSHHEEDADNIVY 106
 QY 136 PFTR-THGDMHSLVTRRHRIPEPEAAVLFROMATALAHCHQGLVLRDILKLFVADSEK 194
 Db 107 FLELCRSKSLAHWKARHILPEPVRYIRQLSLGKYIHQGISILHRDLKNGFITT-- 164
 QY 195 RKKLVLENEDSCVLTPGDDSLMDKAC--PAYVGEIILSPRASYSGKAADVWSLGVLFM 252
 Db 165 NMELKVGDPGLAARLEPEQR--KTIICGTPVVAPEVLLRQG--HGPEADVWSLGCWVY 220
 QY 253 TMLAGHYPRQDSEPVLLFKIRGAYALPAGISAPRCVRCILRREPARELTATGILLH 312
 Db 221 TLGCSPPPETADLETRYCRIKQVHTLPASLSLPARQLAALTRASPDRPSIEQILRH 280
 QY 313 PWLRQ---DPMPLAPTRSHLWEAAQVVPD 338
 Db 281 DFTTGYTPDRLPVS-----SCVYVPD 302

RESULT 10

KKK1 YEAST

ID KKK1_YEAST STANDARD; PRT; 1518 AA.

AC P3424; 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Probable serine/threonine-protein kinase YKL101W (EC 2.7.1.-).
 GN YKL101W OR YKL453.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA MEDLINE=94078677; PubMed=8256524;
 RA Paller C., Valens M., Puzos V., Fukuhara H., Cheret G., Sor F.,
 RA Bolognin-Fukuhara M.;
 RT "DNA sequence analysis of a 17 kb fragment of yeast chromosome XI

```

RT physically localizes the MRB1 gene and reveals eight new open reading
RT frames, including a homologue of the KINI/KIN2 and SNF1 protein
RL Yeast 9:1149-1155(1993).
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIM1
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X71133; CAAS0456.1; -.
DR EMBL, Z28101; CAAB1941.1; -.
DR PIR, S37928; S37928.
DR HSSP, Q63450; 1A06.
DR Germonline; 139857; -.
DR SGD, S0001584; HSIL.
DR GO, GO:0005935; C:bud neck; IDA.
DR GO, GO:0005940; C:septin ring; IDA.
DR GO, GO:0004672; F:protein kinase activity; IDA.
DR GO, GO:0004672; F:G2/M transition of mitotic cell cycle; IGI.
DR GO, GO:0000086; P:G2/M transition of mitotic cell cycle; IGI.
DR GO, GO:0006468; P:protein amino acid phosphorylation; IDA.
DR GO, GO:0000074; P:regulation of cell cycle; IMP.
DR GO, GO:0000135; P:septin checkpoint; IGI.
DR InterPro, IPR000719; Prot kinase.
DR InterPro, IPR008271; Ser_thr_kin_AS.
DR InterPro, IPR002290; Ser_thr_kinase.
DR InterPro, IPR001245; Tyr_kinase.
DR Pfam, PF00069; pkinase; 1.
DR PRINTS, PR00109; TYRKINASE.
DR ProDom, PD000001; Prot kinase; 1.
DR SMART, SM00220; S_TKC; 1.
DR PROSITE, PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE, PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE, PS00111; PROTEIN_KINASE_DOM; 1.
DR Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KM ATP-binding.
FT DOMAIN 81 369 PROTEIN KINASE.
FT NP_BIND 87 95 ATP (BY SIMILARITY).
FT BINDING 110 110 ATP (BY SIMILARITY).
FT ACT_SITE 239 239 BY SIMILARITY.
SQ SEQUENCE 1518 AA; 169592 MW; 803F84F7531241DD CRC64;

Query Match 15.0%; Score 284; DB 1; Length 1518;
Best Local Similarity 26.4%; Pred. No. 4.3e-14;
Matches 85; Conservative 51; Mismatches 128; Indels 58; Gaps 9;

```

```

DB 356 KRITTOELIKHPLIKKYDDLPUV 377
RESULT 11
SNF1_CANGA
ID SNF1_CANGA STANDARD; PRT; 611 AA.
AC Q00372;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
GN SNF1.
OS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCLS84;
RX MEDLINE=97101049; PubMed=8945576;
RA Pether R., Kwon-Chung K.J.;
RT "Disruption of the SNF1 gene abolishes trehalose utilization in the
RT pathogenic yeast Candida glabrata.";
RL Infect. Immun. 64:5269-5273(1996).
CC -1- FUNCTION: Essential for release from glucose repression. It
CC interacts and has functional relationship to the regulatory
CC protein SNF4. Could phosphorylates Ctr8 (By similarity).
CC -1- SUBCELLULAR LOCATION: Associated with the nuclear membrane (By
CC similarity).
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, L78130; AAB48642.1; -.
DR HSSP, P24941; IHCL.
DR InterPro, IPR000719; Prot kinase.
DR InterPro, IPR008271; Ser_thr_kin_AS.
DR InterPro, IPR002290; Ser_thr_kinase.
DR Pfam, PF00069; pkinase; 1.
DR ProDom, PD000001; Prot kinase; 1.
DR SMART, SM00220; S_TKC; 1.
DR PROSITE, PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE, PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE, PS00111; PROTEIN_KINASE_DOM; 1.
DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding;
KM Phosphorylation; Carbohydrate metabolism; Nuclear protein.
FT DOMAIN 6 17 POLY-HIS.
FT NP_BIND 39 230 PROTEIN KINASE.
FT BINDING 45 53 ATP (BY SIMILARITY).
FT ACT_SITE 68 68 ATP (BY SIMILARITY).
FT MOD_RES 161 161 BY SIMILARITY.
FT MOD_RES 194 194 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 611 AA; 70049 MW; 89E17812A4900CD0 CRC64;

Query Match 15.0%; Score 283.5; DB 1; Length 611;
Best Local Similarity 27.9%; Pred. No. 1.6e-14;
Matches 90; Conservative 54; Mismatches 136; Indels 43; Gaps 10;

```

QY 165 QMATALAHCHQHGLVLDLKLCEFEVADREKKVLLENLEDSCVLTGPDSDMKKAC-- 222
 Db 144 QIISAVEYCHRHKIVHDLKPEMLLDEHNLVXKIDGSL--SNMT---DSEPLKTSGS 198
 QY 223 PAYVGPILSRASYSGKADVMSLGVALLFTMLAGHPPODSEVLLFGKIRGAYALPA 282
 Db 199 PNVAPEVITSKL-YAGEVDVWMSGVLLVYMLCRRLPFDEESIPVLFKINSNGVYTLPK 257
 QY 283 GLSAPARCLVRCILRRPAPERLTATGILLHPWLRQDPMPLAPFTSHLMEAAQVYDDGL-- 340
 Db 258 FLISGASDLIKRMILVNPARKISHEIMODEMFVYD-----LAEYLVPODLKQ 305
 QY 341 -----GLDEAREEGSDREV 355
 Db 306 QEOFNKKSQNEENVEIDDEWV 328

RESULT 12
 ID K110_ARATH STANDARD; PRT; 535 AA.
 AC Q38997; 004728; Q39076; Q8RWD2;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-MAR-2004 (Rel. 43, last sequence update)
 DE SNRP1-related protein kinase KIN10 (EC 2.7.1.1) (AKIN10).
 GN KIN10 OR SKIN10 OR AT3G01090 OR T4P13.22.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopses.
 NC NCB1_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=93013041; PubMed=1339373;
 RA 1e Guen L., Thomas M., Bianchi M., Halford N.G., Kreis M.;
 RT "Structure and expression of a gene from Arabidopsis thaliana
 encoding a protein related to SNF1 protein kinase.";
 RL Gene 120:249-254(1992).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX STRAIN=cv. Columbia;
 RA Lessard P., Kreis M., Thomas M.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. Columbia;
 MEDLINE=21016720; PubMed=11130713;
 RA Salanoubat M., Lemcke K., Rieger W., Ansoerge W., Unseld M.,
 RA Farreman B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
 RA Delany M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
 RA De Simone V., Choisy N., Artiguenave F., Robert C., Brotier P.,
 RA Wincker P., Catolico L., Weissenbach J., Saurin W., Queller F.,
 RA Schaefer M., Meller-Auer S., Gabel C., Fuchs M., Benes V.,
 RA Wurmbach E., Drzonek H., Erle H., Jordan N., Bangert S.,
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brand P., Nyakatura G.,
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simionati B.,
 RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordsiek G.,
 RA Reichelt U., Schafte M., Schoen O., Barques M., Teol J., Clement J.,
 RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
 RA Cooke R., Lande M., Berger-Liauro C., Purnelle B., Masny D.,
 RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottec R., Casasuberta E.,
 RA Monfort A., Argitron A., Flores M., Liguori R., Vitale D.,
 RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T., Rizzo M., Walts A., Utecherback T., Fujii C.Y., Shea T.P.,
 RA Cressy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
 RA Pai G., Miltischer J., Sellers P., Gill J.E., Feldlyum T.V.,
 RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
 RA Frazer C.M., Kaneo T., Nakamura Y., Sato S., Kato T., Asamizu E.,
 RA Sasamoto S., Kimura T., Ideawa K., Kawashina K., Kishida Y.,
 RA Kiyokawa S., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;

RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:820-822(2000).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX STRAIN=cv. Columbia;
 MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Kim U., Dale J.M., Chen H., Shin P., Palm C.J.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Atiyana K., Aneri Y.,
 RA Arakawa T., Ban J., Banno F., Bowser L., Brooks S.Y., Carnici P.,
 RA Chao O., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tame R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 genome.";
 RL Science 302:842-846(2003).
 RN [5]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM 2).
 RX STRAIN=cv. Columbia;
 MEDLINE=95115691; PubMed=7816049;
 RA 1e Guen L., Thomas M., Kreis M.;
 RT "Gene density and organization in a small region of the Arabidopsis
 thaliana genome.";
 RL Mol. Gen. Genet. 245:390-396(1994).
 CC -1- FUNCTION: May play an important role in a signal transduction
 cascade regulating gene expression and carbohydrate metabolism in
 higher plants.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q38997-1; Sequence=displayed;
 CC Note=No experimental confirmation available;
 CC Name=2;
 CC IsoId=Q38997-2; Sequence=VSP_009001;
 CC -1- TISSUE SPECIFICITY: Expressed in roots, shoots and leaves.
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNRP1
 subfamily.
 CC -1- SIMILARITY: Contains 1 UBA domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib.ch).
 CC -----
 DR EMBL; M93023; AAA2736.1; -;
 DR EMBL; X79707; CAA56146.1; -;
 DR EMBL; AC008261; AAF26165.1; -;
 DR EMBL; AY093170; AAM3169.1; -;
 DR EMBL; BF010386; AA056829.1; -;
 DR EMBL; X94757; CAA64384.1; -;
 DR FPIR; JCI1446; JCI1446.
 DR HSP; O63450; 1A06.
 DR InterPro; IPR001712; Kinase Cterm.
 DR InterPro; IPR000719; Prot Kinase.
 DR InterPro; IPR008271; Ser_Thr_kin_AS.
 DR InterPro; IPR002290; Ser_Thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF02149; KAI; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00627; UBA; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SMO0220; S_TKc; 1.

DR SMART; SM00165; UBA; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Alternative splicing.
 FT DOMAIN 42 294
 FT NP BIND 48 56 ATP (BY SIMILARITY).
 FT DOMAIN 315 355 UBA.
 FT ACT SITE 165 165 BY SIMILARITY.
 FT BINDING 71 71 ATP (BY SIMILARITY).
 FT MOD RES 198 198 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT VARSPLIC 1 23 Missing (in isoform 2).
 FT /FTID=VSP 009001.
 SQ SEQUENCE 535 AA; 61181 MW; FPC038322FDB317 CRC64;
 Query Match 14.8%; Score 280; DB 1; Length 535;
 Best Local Similarity 28.7%; Pred. No. 2.6e-14;
 Matches 87; Conservative 46; Mismatches 118; Indels 52; Gaps 9;
 QY 80 GRAYQALHCPGTCTCKVY-----PVQALAVLEPYRLPHKHAVPTETL 127
 DB 54 GRVXIAEHALTGKVAIKILNRKIKNMMEKVRREIKIL---RLFMHPHILRYEVI 109
 QY 128 AGTOLLVAFPTRTGDMHSLVTRRIPEPEAAVLFROMATALACHQGLVLRDLK-- 184
 DB 110 EPPFTDIYVMEYVNSSELPDYIVEKRLQEDEARNFQQLISVEGCHNNVVRDLKDE 169
 QY 185 -----LCRFVADRRERKVLLEN--LEDSCVLGTGDDSLMDKHCAPVYGEIISRRAS 236
 DB 170 NLLDSKCNVKIADFLSLNMRDGHPLKTSC-----GSPNVAABEVISGKL- 215
 QY 237 YGSKADVMSLGVAFPTMAGHYPPQDSEPVLLFGKIRGAYALPAGLSAPRCVRLCL 296
 DB 216 YAGPEVDVMSGCIYVALCGTLPFDENIPNLFKKIKGCIYLPBHLSPGARDLIPRWL 275
 QY 297 RRPAPRLTATGTLHPWLRQDMPPLAPTRSHLMEAAQV--VPDGLDLEAREEGRREV 355
 DB 276 VVDPMKRVITLPIRQHPW-----QAHLPRLVAPPPPTVQAKKIDELIGEV 325
 QY 356 IYVG 358
 DB 326 NMVG 328
 RESULT 13
 SNF1_YEAST STANDARD; PRT; 633 AA.
 ID SNF1_YEAST
 AC P06782;
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-JAN-1988 (Rel. 06, Last annotation update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update) (EC 2.7.1.-)
 DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-)
 GN SNF1 OR CAT1 OR CAS14 OR GUC2 OR YDR477W OR DB035.20.
 OS *Saccharomyces cerevisiae* (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OC NCBI_TaxID=4932;
 OX [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=86289463; PubMed=3526554;
 RA Celenza J.L., Carlson M.;
 RT "A yeast gene that is essential for release from glucose repression
 RT encodes a protein kinase."
 RL Science 233:1175-1180 (1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
 RP Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
 RP Hunnicke-Smith S., Hyman R., Komp C., Lahekari D., Lew H., Lin D.,
 RP Moseedale D., Nakahara K., Namath A., Oetner P., Or C., Fétel F.X.,
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,

RA Winant A., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 274-284; 528-539 AND 622-630, AND PHOSPHORYLATION SITE.
 RX MEDLINE=94131988; PubMed=7905477;
 RA Mitchellhill K.I., Stapleton D., Gao G., House C., Mitchell B.,
 RA Kassis F., Walters U.A., Kemp B.R.;
 RT "Mammalian AMP-activated protein kinase shares structural and
 RT functional homology with the catalytic domain of yeast Snf1 protein
 RT kinase."
 RL J. Biol. Chem. 269:2364-2364 (1994).
 CC -FUNCTION: Essential for release from glucose repression. It
 CC interacts and has functional relationship to the regulatory
 CC protein SNF4. Interacts also with S1P1, S1P2 and GAL83. Could
 CC phosphorylates CAT8.
 CC -SUBCELLULAR LOCATION: Associated with the nuclear membrane.
 CC -SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
 CC subfamily.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL; M13971; AAA35058.1; -;
 CC EMBL; U33050; AAB64904.1; -;
 CC PIR; A26030; A26030.
 CC HSRP; P24941; 1HCL.
 CC Germonline; 140969; -;
 CC SGD; S0002885; SNF1.
 CC GO; GO:0005737; C:cytoplasm; IPI.
 CC GO; GO:0005634; C:nucleus; IPI.
 CC GO; GO:000324; C:cytosol (sensu Fungi); IPI.
 CC GO; GO:0004679; P:SNF1A/AMP-activated protein kinase activity; IDA.
 CC GO; GO:0007155; P:cell adhesion; IMP.
 CC GO; GO:0006995; P:cellular response to nitrogen starvation; IDA.
 CC GO; GO:0030447; P:filamentous growth; IMP.
 CC GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
 CC GO; GO:0006109; P:regulation of carbohydrate metabolism; IGI.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR008271; Ser_Thr_kin_AS.
 CC InterPro; IPR002290; Ser_Thr_kinase.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot_kinase; 1.
 CC SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Carbohydrate metabolism; Nuclear protein.
 FT DOMAIN 18 32
 FT NP BIND 55 306 PROTEIN KINASE.
 FT BINDING 61 69 ATP (BY SIMILARITY).
 FT ACT SITE 84 84 ATP (BY SIMILARITY).
 FT MOD RES 177 177 BY SIMILARITY.
 FT VARSPLIC 210 210 PHOSPHORYLATION (AUTO-).
 FT /FTID=VSP 009001.
 SQ SEQUENCE 633 AA; 72045 MW; F5C63565C986C4E3 CRC64;
 Query Match 14.5%; Score 273.5; DB 1; Length 633;
 Best Local Similarity 27.8%; Pred. No. 1e-13;
 Matches 88; Conservative 53; Mismatches 138; Indels 37; Gaps 11;
 QY 51 SPTAPDARATVASHRGVYLLER--EEG--GRAYQALHCPGTCTCKVYVQALAV 107
 DB 35 SNTSLNPKSSLDGAGHAGNYQIVKTGSGSFGVKYLAHTTTGQXALMLN-KYVLAK 93
 QY 108 LE-----PYARLPFHGVAPTEVLVAGTQLLVAFRTTGHGMSLVTRHRIPEPE 158
 DB 94 SDWGRIERERISTVRLRHRIHIIKLVDVSKDEIIVIVAGNELFDYIVQDRKSGOE 153

```

QY 159 AAVLFQWATALAHCHQHLVRLDKLRFVAFADREKVLLENLDESCVLTGPDDSLMD 218
DQ 154 ARRFQOQIISAVEYCHRHKIVHRLDKPENLLDHLNWKADPGL--SNIT--DGNFL 208
QY 219 KHAC--PAYQPELLSPASTSGAADVSTIGVALFTMLAGHYPODESEPVLPFGKTRG 276
DQ 209 KTSQSGSPYAPAEVISGEL-YAGPEVDVWSCGVILLYWMLCRRLPFDDESIIVLPKINSIG 267
QY 277 AVALPAGISAPARCIVCLIFEREAPERLTATGILLHMLRPD--PMPLAPTRSHLMEAAQV 335
DQ 268 VYTLPKLSPGAAAGIKRMILVNPINRISIHETWDDMFVDPLEVILP----- 316
QY 336 VPDGLGDEAREEBCD 351
DQ 317 -PD---LKPHPEENE 328

RESULT 14
CDRL_SCHPO STANDARD; PRT; 593 AA.
ID CDRL_SCHPO STANDARD; PRT; 593 AA.
AC P07334; Q9PEQ4;
DT 01-APR-1988 (Rel. 07, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mitosis inducer protein kinase cdrl (EC 2.7.1.-) (Protein kinase
DE nm1).
GN CDRL OR NIM1 OR SPAC644.06C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91169281; PubMed=2004705;
RA Fellotter H., Nurse P., Young P.G.;
RT "Genetic and molecular analysis of cdrl/nim1 in Schizosaccharomycetes
RT pombe.";
RL Genetics 127:309-318 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87187654; PubMed=3453113;
RA Russell P., Nurse P.;
RT "The mitotic inducer nim1+ functions in a regulatory network of
RT protein kinase homologs controlling the initiation of mitosis.";
RL Cell 49:569-576 (1987).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Baaham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Frazer A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones I., Jones M., Leather S., McDonald S., McLean T.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymoprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mellertner S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moser D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Coffeau A., Gadieu E., Dreno S., Gloux S., Lelaure V., Mottier S.,
RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Ruvetta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

```

```

RA Shpakovski G.V., Useery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880 (2002).
CC -1- FUNCTION: This protein, a dose-dependent mitotic inducer, appears
CC to function as a negative regulator of mitosis inhibitor weel by
CC phosphorylating and inactivating it.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIM1
CC subfamily.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to
CC frameshifts.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X57549; CAA40774.1; -.
CC EMBL; M16509; AAA35317.1; ALT_FRAME.
CC EMBL; ALJ55012; CAB90133.1; -.
CC PIR; S16153; KIZPMN.
CC HSSP; O63450; 1A06.
CC GeneDB Spombe; SPAC644.06C; -.
CC InterPro; IPR000719; Prot_Kinase.
CC InterPro; IPR008271; Ser_Thr_pkin_AS.
CC InterPro; IPR002290; Ser_Thr_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_Kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC KMW Transferrase; Serine/chreonine-protein kinase; ATP-binding;
CC Mitosis.
CC KW
CC FT DOMAIN 12 258 PROTEIN KINASE.
CC FT NP_BIND 18 26 ATP (BY SIMILARITY).
CC FT BINDING 41 41 ATP (BY SIMILARITY).
CC FT ACT_SITE 128 128 BY SIMILARITY.
CC FT CONFLICT 252 252 V -> F (IN REF. 1).
CC FT CONFLICT 570 570 A -> I (IN REF. 1).
CC SQ SEQUENCE 593 AA; 66954 MW; DE30AE06B070F458 CRC64;

```

Query Match 14.3%; Score 270.5; DB 1; Length 593;
 Best Local Similarity 31.5%; Pred. No. 1.6e-13;
 Matches 74; Conservative 45; Mismatches 107; Indels 9; Gaps 6;

```

QY 85 AHCPTGTEYTKVYPVOALAVD--PYARLPKHVAPTEVLATGTOILY-AFPTRTG 142
DQ 29 AHAATGDLAAKIKIPRYASIGMEILMMRLRHNPILRLYVDWHDQHMVLALEYVPG 88
QY 143 DMSHIVTRTRRIPREPAVLFROMATLALAHCHQHLVRLDKLRFVAFADREKVLLEN 202
DQ 89 ELFHTRKRGPSERBAHYLSQILDVAHCHRPFRRHRLDLENILIKVNDQ--IKI 145
QY 203 LEDSCVLTGPDDSLMDKACPA--YVGPBILSSRASYSKADVWSLGVALTMLAGHYP 260
DQ 146 ADFGAWATEPNDSCLENY-CGSIHLTALEIVSHK-PYGAAPADWSCGVILYSLSNKLP 203
QY 261 FQDSFVLLFNGTRGATLPAAGLSAPARCIVCLIFEREAPERLTATGILLHMLPWT 315
DQ 204 FQGQNTDVIYNKIRGAVDLPSSISSAODILHRLMDVNPSTRITRIVFVSHPL 258

```

```

RESULT 15
SNP1_SCHPO STANDARD; PRT; 576 AA.
ID SNP1_SCHPO STANDARD; PRT; 576 AA.
AC Q074536;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)

```

DT 10-OCT-2003 (Rel. 42, last annotation update)
 DE SNF1-like protein kinase (BC 2.7.1.-).
 GN SPC74.03C.
 OS Schizosaccharomyces pombe (Pisision Yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Galliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Pat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean U.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben U., Grymoprez B.,
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wandt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Delaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez J., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomyces pombe."
 RL Nature 415:871-880(2002).
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
 subfamily.
 CC -!- SIMILARITY: Contains 1 UBA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL031543; CAA20833.1; -;
 DR PTR: T41587; T41587.
 DR HSSP: Q63450; 1A06.
 DR GeneDB: SPombe, SPC74.03C; -;
 DR InterPro: IPR000719; Ser_kinase.
 DR InterPro: IPR008271; Ser_thr_kin AS.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR000449; UBA_domain.
 DR Pfam: PF00069; kinase_1.
 DR ProDom: PD000001; Prot_Kinase; 1.
 DR SMART: SMO0220; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50030; UBA_1.
 DR Transferase: Serine/threonine-protein kinase, ATP-binding.
 KW TRANSFERASE; Serine/threonine-protein kinase, ATP-binding.
 FT DOMAIN 34 285 PROTEIN_KINASE.
 FT NP_BIND 40 48 UBA.
 FT BINDING 63 63 ATP (BY SIMILARITY).
 FT ACT_SITE 156 156 ATP (BY SIMILARITY).
 FT ACT_SITE 156 156 BY SIMILARITY.
 SQ SEQUENCE 576 AA; 65996 MW; E5857E8F171E7B50 CRC64;

Query Match 14.3%; Score 270; DB 1; Length 576;
 Best Local Similarity 25.8%; Pred. No. 1.7e-13;
 Matches 83; Conservative 54; Mismatches 133; Indels 52; Gaps 11;
 QY PPTAPDRATAVATASRLGPVYLLEP-BEG--GRAYQALHCPGTGEYTKVYP----- 100
 DB 22 PPEA-----ISKRHITPIIIRFTLGSQKVLALHYTKQKVALKFTISRLKSSD 74
 QY 101 ---VQELAVLEPYARLPKPKHVARPREVLATGQLLYAFPTRTGDMESLVTRTHRIPE 156
 DB 75 MEMRVEREIS---YTKLRHPHITIKLDVITTPDVIWVLEIYAGGELFDYIVKKRMTE 130
 QY 157 PEAAVLFPQMATALAHQHOGIVLBDLKLCPVFPADREKKIVLENLEDSCVLTGPDSTL 216
 DB 131 DEGRFPFOOICALIYCHRHKIVHDLKPENLIDNNVKKIADFGI--SNIT--DGN 185
 QY 217 WDKHAC--PAYVPELTSSRASYSKADVMSLGVAFPTMLAGHYFPQDSPEVLLFGKIR 274
 DB 166 FLKTCGSPNVAAPVINGKL-YAGPEVDVMSGGLVLYVMLVGRLPFDDEFIPNLFKYN 244
 QY 275 RGAVYLPAGLSAPACLVRCCLRRPAPRLTATGILLHPWRQDPMPLAPTRSHMEAAQ 334
 DB 245 SCVYVMPDFLSPQASLIRMTIVADPMQRIITQETIRDPWFNVN----- 288
 QY 335 VVPDGL-GIDKAREEGPREV 355
 DB 289 -LPDYLRPMEEVQSGSYADSRIV 309

Search completed: August 24, 2004, 18:53:20
 Job time : 14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 18:49:06 ; Search time 45 Seconds
(without alignments)
2510.123 Million cell updates/sec

Title: US-10-070-337-5
Perfect score: 1892
Sequence: 1 MRATPLAAPAGSLSRKRL.....GLGLDEAREBEGREVLYG 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: SP archaea:*
 - 2: SP bacteria:*
 - 3: SP fungi:*
 - 4: SP human:*
 - 5: SP_invertebrate:*
 - 6: SP_mammal:*
 - 7: SP_mhc:*
 - 8: SP_Organelle:*
 - 9: SP_phase:*
 - 10: SP_plant:*
 - 11: SP_rodent:*
 - 12: SP_virus:*
 - 13: SP Vertebrate:*
 - 14: SP Unclassified:*
 - 15: SP Virus:*
 - 16: SP_bacteriap:*
 - 17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	Length	DB ID	Description
1	766.5	40.5	343	13	07ZYZ2
2	759.5	40.1	343	6	0864R4
3	756	40.0	343	11	08K017
4	754.5	39.9	343	6	028283
5	753.5	39.8	343	4	092519
6	753	39.8	343	11	08K4K3
7	742.5	39.2	364	11	09EG16
8	737.5	39.0	372	4	096R18
9	735.5	38.9	372	4	09H218
10	735.5	38.9	372	11	091W04
11	722.5	38.2	372	11	08K4K4
12	717	37.9	344	13	07Z2Z7
13	657	34.7	257	11	08R2V8
14	608.5	32.2	224	4	015180
15	432	22.8	484	5	09V3Z1
16	341.5	18.0	443	10	07X1W5

17	319	16.9	218	11	08BFS7	08BFS7 mus musculu
18	319	16.9	379	11	08BUR9	08BUR9 mus musculu
19	319	16.9	388	11	08BZX3	08BZX3 mus musculu
20	318	16.8	461	10	09LGV5	09LGV5 oryza sativ
21	313.5	16.6	706	3	08J2N0	08J2N0 fusarium ox
22	305	16.1	474	10	094D77	094D77 oryza sativ
23	303.5	16.0	446	4	096CV1	096CV1 homo sapien
24	299	15.8	914	5	019469	019469 caenorhabdi
25	298.5	15.8	446	10	08H5S1	08H5S1 oryza sativ
26	298	15.8	504	11	08K0J7	08K0J7 mus musculu
27	296	15.6	880	3	09Y880	09Y880 cocholebolu
28	295	15.6	671	3	096M17	096M17 trichoderma
29	293	15.5	777	3	0872H0	0872H0 neurospora
30	291.5	15.4	443	13	07T0B0	07T0B0 xenopus lae
31	291.5	15.4	651	13	091821	091821 xenopus lae
32	290	15.3	602	3	P87209	P87209 kluyveromyc
33	288.5	15.2	698	5	061298	061298 halocynthia
34	288	15.2	688	5	095U14	095U14 ancylostoma
35	287.5	15.2	676	13	07ZZN5	07ZZN5 brachydanto
36	287	15.2	477	10	09LW16	09LW16 oryza sativ
37	286.5	15.1	676	13	07ZU72	07ZU72 brachydanto
38	284	15.0	514	10	09M726	09M726 lycopersico
39	283	15.0	504	10	P93113	P93113 cucumis sat
40	283	15.0	512	10	P92958	P92958 arabidopsis
41	282	14.9	438	10	08LX24	08LX24 glycine max
42	282	14.9	512	10	P92968	P92968 arabidopsis
43	281.5	14.9	419	4	09H7H6	09H7H6 homo sapien
44	280.5	14.8	643	11	061804	061804 mus musculu
45	280.5	14.8	643	11	061846	061846 mus musculu

ALIGNMENTS

RESULT 1

07ZYZ2 ID 07ZYZ2 PRELIMINARY; PRT; 343 AA.
AC 07ZYZ2;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE TRB2 protein.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Shan Y.X., Huang C.Q., Yu L.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY247742; AAP0411.1; -
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.
SQ SEQUENCE 343 AA; 38875 MW; C87A719FB016B42D0 CRC64;

Query Match 40.5%; Score 766.5; DB 13; Length 343;

Best Local Similarity 47.6%; Pred. No. 2.9e-57;
Matches 169; Conservative 50; Mismatches 114; Indels 21; Gaps 6;

QY 2 RAEPLAAPAGSLSRKRLLELDNLDTERPVOKRARGPQPRLLPCLLPPTAPAPRATA 61
DB 5 RSPDITARYGRGRNKTQDEE-LSSIRSLD-----PQSGFSPNUGSPDPPEPPNLSHC 57


```

QY 179 VRDRLKLGCFVADNRKKTLVLENNEDSCVLCTPDSDLMWDKACPAVYGEILLSPRASY 238
Db 172 VRDRLKLRFRKDEERTVVKLESTEDAYILRRDDDSLSPKHCPAAVSEILLNTSGSYS 231
QY 239 GKAAVWMSGVALFTMLAGHYPPQDSEPVLLFGKIRGAYALPAGLSAPRCLVRLTLR 238
Db 232 GKAAVWMSGVALFTMLVGRYPHPHDLPSSLFKIRGGQNPETLSPKAKCLIRSLTLR 231
QY 259 EEAERLTATGILLHPWLRDPMPLAFTRSHLWMAAQVPEGLDLEAREE 348
Db 252 EEESEKLTSEILLDHPWFSTD-----FVSVNSGFGAKACACDQ 327

```

ID	Q28283	PRELIMINARY;	PRT;	343 AA.
AC	Q28283;			
DT	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-NOV-1996 (TrEMBLrel. 01, last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, last annotation update)			
DE	C5FW ORF protein.			
GN	C5FW ORF.			
OS	Canis familiaris (Dog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
OX	NCBI_TaxId=9615;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Thyroid.			
RX	MEDLINE=97067069; PubMed=8910471;			
RA	Wilkin F., Savonet V., Radulescu A., Petermans J., Dumont J.E.,			
RA	Maenhaut C.;			
RT	"Identification and Characterization of Novel Genes Modulated in the			
RT	Thyroid of Dogs Treated with Methimazole and Propylthiouracil.";			
RL	J. Biol. Chem. 271:28451-28457(1996).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Thyroid.			
RX	MEDLINE=98000262; PubMed=9342215;			
RA	Wilkin F., Suarez-Huerta N., Robaye B., Peetermans J., Libert F.,			
RA	Dumont J.E., Maenhaut C.;			
RT	"Characterization of a phosphoprotein whose mRNA is regulated by the			
RT	mitogenic pathways in dog thyroid cells.";			
RL	Eur. J. Biochem. 248:660-669(1997).			
DR	EMBL; X99144; CAA67581.1; -.			
DR	HSSP; Q63450; 1A06.			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0004672; F:protein kinase activity; IEA.			
DR	GO; GO:0016740; F:transferase activity; IEA.			
DR	GO; GO:0004668; F:protein amino acid phosphorylation; IEA.			
DR	InterPro; IPR000719; Prot_Kinase.			
DR	Pfam; PF00069; pkinase; 1.			
DR	ProDom; PD000001; Prot kinase; 1.			
DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.			
KW	ATP-binding; Transferase.			
SQ	SEQUENCE 343 AA; 38786 MW; BF8D1300DACB84FA CRC64;			
QY	Query Match	39.9%;	Score 754.5;	DB 6; Length 343;
	Best Local Similarity	47.3%;	Pred. No.3e-56;	
	Matches 167; Conservative	50;	Mismatches 115;	Indels 21; Gaps 6
DB	2 RATPLAAPAGSGISRRKRLDLDDLDLDERPVQKARASGPQRRLPCLLSPPTAPADATA	61		
QY	62 VATAASRLGPIVLLPEPEEGRAYOALCPTGTEYTKCKVAPV---QEALAVLEPYARLPPIK	118		
DB	58 V---SGTGYLLLEPEGDHVFRAVHLHSGEELVCKVFDSICQGESLA---PCPCSAHS	111		
QY	119 HVAAPPEVLAAGTOLLAFTPTRTGDMGSLVTRTHRIPEPAALVFROMATALAHCHQGL	178		
DB	112 NNQQLTEIILGETKAVVFFERSGDMSFVTKCKKKEEENARLFYQIASAVACHDGGT	171		

```

QY 179 VLRLDKLRFPADEERKKVLVLEINEDSOVLVLPGRDLSMDKKACPAVYVEPELISASBAS 238
Db 172 VLRLDKLRFKPFKEERTVRKTESLEDAVILFLGDSDSLDKKGCPAVYVEPELINTGSYS 231
QY 239 GKAAVWSLGLVLFMLAGHYFPDSEPFVLFGKIRGAYALPAGISAPARCCLVRLLR 298
Db 232 GKAAVWSLGLVLYMLVGRYPFHIEBSLFSKIRRGQFENLPETISPAKCLITRSLRR 291
QY 239 EPARLRTGLILHPMLRODPEMLAPTRSHLWEAAQVVPDGLGLDEAREEGD 351
Db 232 EPERLRTQGEIILHPWFSTD-----FVSVNSGTAKEVSDQLVLPDVMMENL 339

```

Query	Match	Score	753.5	DB 4	Length	343
Best Local Similarity	47.3%	Pred. No. 3.7e-56				
Matches	167	Conservative	50	Mismatches	115	Indels 21; Gaps 6
QY	2 RATPLAAPAGSLRRKKRLDNDLTERPVOKRARGPOPLRPCLLPSPTAPDRATA	61				
DB	5 RSTPITARIYGRSRNKTQDFEE-TLSIRGAE-----PQSFSFNTGSPSPPTPNLSHC	57				
QY	62 VATAARLGGVYLLEPEEGGRAYQALHCPRTGYTCVYVY---QELAVLEFARLPRPK	118				
DB	58 V---SCIGKYLLLEPLGGDHVRAVHLHSGEELVCKVPISCYQESIA--PCFCLSAHS	111				
QY	119 HVARTEVLVAGTQLYAFRTTHGDMHSVLRTRHRIPEEAAVLFRQMATALACHQHGL	178				
DB	112 NINQTEITILGSTKXAVFFPERYGDHMSFVRCKKLREEARLFLQINSAVAHCHDGL	171				
QY	179 VIRDIKLCFVPADERRKKLVIENTEDSCVLITGPDDSLMDKIACPAYVGGPEILISSRAYS	238				

```
DB 172 VLKDLKLRKFIKDEERTVYKLESLDPAVILRGDDSLSDKGGCPAVVSEIINTSGS231
QY 239 GKADVWSLGVALLFTMLAGHYPPQDSEPVLLFGKIRGAYALPAGLSAPRCCLVRLRR 238
DB 232 GKADVWSLGVWLYTLVGRYPFHDIPESSLSFKIRGGQFNIPETLSPPKACILRSLRR 231
QY 299 EPAERLTATGILHLPMLRQDPMPLAFTRSHLWEAAQVDPDGLGDEAREEGD 351
DB 292 EPERLTSOELIDHPWFSTD-----FSVNSAIGAKEVSLQVLPDVAMEEMLD 339

RESULT 6
Q8K4K3 PRELIMINARY; PRT; 343 AA.
AC 08K4K3;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DE 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE TRB-2.
GN AM319517.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RA Kissa-Tsch E., Dempsey C., Jozsa V., Caunt J., Oxley K.M.,
RA Bagnsaff S.M., Wyllie D.H., Harte W., O'Neill L.A.U., Qvarnstrom E.B.,
RA Dower S.K.;
RA "Mammalian homologs of Drosophila tribbles (trib) control mitogen
RT activated protein kinase signaling."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF358867; AA54577.1; -.
DR MCD; MGJ:2145021; AM319517.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR ATP-binding; Transferase.
SQ SEQUENCE 343 AA; 38758 MW; 0B3965B8B2087D74 CRC64;

Query Match 39.8%; Score 753; DB 11; Length 343;
Best Local Similarity 46.6%; Pred. No. 4.1e-56;
Matches 163; Conservative 49; Mismatches 108; Indels 30; Gaps 6;

QY 2 RATELAAPAGSLSRKKRLDNLDERVOKRARGPOPLPCLPLSPPTAPRATA 61
DB 5 RSTPLIARYGSRNKTQDFEE-ISSIRAE-----PSQSPNNGSPSPETPLSLHC 57
QY 62 VATASLGGVYLLEPEGGRAYOALHCPGTETCTCYVYV---QELALVLEPYARLPKH 118
DB 58 V---SCIGKYLLLEPLEGGHVFRAVHLSGEEIVCVFELISCSQESIA--PCFCISAS 111
QY 119 HVAPREVLATGTLAFTTRTHGMSLVTRHRIPEPAVLFPQMTALAHCHQHL 178
DB 112 NINQITRIILGELKAVYFFERSIGDMSVYRTCKLRSEBARLFYQISAAVHCHDGL 171
QY 179 VLKDLKLRKFIKDEERTVYKLESLDPAVILRGDDSLSDKGGCPAVVSEIINTSGS 238
DB 172 VLKDLKLRKFIKDEERTVYKLESLDPAVILRGDDSLSDKGGCPAVVSEIINTSGS 231
QY 239 GKADVWSLGVALLFTMLAGHYPPQDSEPVLLFGKIRGAYALPAGLSAPRCCLVRLRR 298
DB 232 GKADVWSLGVWLYTLVGRYPFHDIPESSLSFKIRGGQFNIPETLSPPKACILRSLRR 291
```

```
QY 299 EPAERLTATGILHLPMLRQDPMPLAFTRSHLWEAAQVDPDGLGDEAREE 348
DB 292 EPERLTSOELIDHPWFSTD-----FSVNSGFGAKACQD 327

RESULT 7
Q9EQL6 PRELIMINARY; PRT; 364 AA.
AC 09EQL6;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE G-protein-coupled receptor induced protein G1G2 (Fragment).
GN G1G2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Mayhaus M., von der Kammer H., Klaudiny J., Albrecht C., Hoffmann B.,
RA Nitsch R.M.;
RT "Identification of a novel nuclear factor G1G2, as an ml-acetylcholine
RT receptor-induced gene."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF205438; AAG3664.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; pkinase; 2.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Receptor; Transferase.
FT NON_TER 1
SQ SEQUENCE 364 AA; 40377 MW; 8F9167B76D76D37 CRC64;

Query Match 39.2%; Score 742.5; DB 11; Length 364;
Best Local Similarity 45.6%; Pred. No. 3.5e-55;
Matches 161; Conservative 48; Mismatches 119; Indels 25; Gaps 5;

QY 2 RATELAAPAGSLSRKKRLDNLDERVOKRARGPOPLPCLPLSPPTAP-- 56
DB 10 RSPALFFPARAGTPAKRL-----LDTDDAAVAAKC---PRLSECSNPPDYLSPPGSPCS 61
QY 57 -----DRATVATASRLGPVYLLEPEGGRAYOALHCPGTETCTCYVYQEL 105
DB 62 POPPSAOGTGS CVSAPGPSRIADTLPLAERHVSALCIHTGERCKVFPKHYQ 121
QY 106 AVLEPYARLPKHVARPTEVLATGTLAFTTRTHGMSLVTRHRIPEPAVLFPQ 165
DB 122 DKIRPYTQLPSRNRITGIVELIGESKAVYFEKPGDMSVYRSKRLREBARLFQ 181
QY 166 MATALAHCHQHLVLDLKCRFVADRERKLVLENLSDCVLTGPDSDLMDKACPAY 225
DB 182 IYSAVAHCHQSAIVIGDGLRKRFVSTERTQLRLESLEDITMIKEDALSDKHCCPAY 241
QY 226 VGEPIIISPRASYSGKAADVWSLGVALLFTMLAGHYPPQDSEPVLLFGKIRGAYALPAGS 285
DB 242 VSEPIINTTGYSGKAADVWSLGMVLYTLVGRYPHSDPSALFSKIRGGQFCLEHVS 301
QY 286 APARCLVRLRRERPAERLTATGILHLPMLRQDPMPLAFTRSHLWEAAQVDP 338
DB 302 PARACILRSLRRERPAERLTATGILHLPWFVEYVLE--GYVDSVGYGSDIIVE 353

RESULT 8
Q96RUB PRELIMINARY; PRT; 372 AA.
ID 096RUB
```


AC Q96RU8; (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (T-EMBLrel. 25, Last annotation update)
 DE SKIPL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kiss-Toch E., Wylie D.H., Qvarnstrom E.E., Dower S.K.;
 RT "Identification of pro-inflammatory cytokine signalling network
 components by transcription expression screening."
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF250310; AAKS8174.1; -
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; pkinase; 2.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR ATP-binding; Transferase.
 KW SEQUENCE 372 AA; 40980 MW; 3E2B5C87A4F98FDB CRC64;
 SQ

Query Match 39.0%; Score 737.5; DB 4; Length 372;
 Best Local Similarity 46.2%; Pred. No. 9.5e-55;
 Matches 160; Conservative 49; Mismatches 126; Indels 11; Gaps 6;

QY 2 RATPLAAPA-GSLSRKKRLLEDD--NLDTERPVQKRAASGP---QRLPFCLLPSPPA 55
 DB 17 RGPALLFPATKRCVPKRLDADDAVAAKCPRLSECCSPDYLSPPGSPC-SPOPPPA 75
 QY 56 PDRATAVATA--SRLGPFVLLPEPEGGRAYQALHCPGTETXCKVYPQALAVLEPYA 112
 DB 76 PGAGGSGSAPSPRSIADYLLPLAEREHVSALCIHGRRLRCVFPPIKHQDKIRPYI 135
 QY 113 RLPRKHVARPTEVLATGQLVAFTRTHGDMHSLVTRHRIPEPEAAVLFROMATALAH 172
 DB 136 QLPSSHSNITGIVEVILGETKAYVFPFKSGDMHYSVRSKRLREBEAARLFQIYSAVAH 195
 QY 173 CHQHGLVLRDLKLCGFVADDERKTLVLENDSCVLTPGPDLSLMDKACPAVYGPETLS 232
 DB 196 CHQSAIVIGDLKRFVFTSTERTQLRLESLEDTIMKGEDDALSCKGCPAYVSPETLN 255
 QY 233 SRASYSGRADAVMSLGVALLFTMLAGHYPFODESEPVLLFGKIRGAYALPAGISAPARCLV 292
 DB 256 TTGTYSGRADAVMSLGVALLFTMLAGHYPFODESEPVLLFGKIRGAYALPAGISAPARCLV 315
 QY 293 RCLLRPEPAERLTATGILLHFWLRQDPMPLAFTSRSHLMEAAQVVD 338
 DB 316 RSLLRPEPSERLTAPEILLHPFESVLEP-GYIDSEIGTSQIYVE 360

RESULT 9

Q9H2Y8 PRELIMINARY; PRT; 372 AA.
 AC Q9H2Y8;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE G-protein-coupled receptor induced protein G1G2.
 GN G1G2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Mayhans M., von der Kammer H., Klaudiny J., Albrecht C., Hoffmann B.,

RA Nitsch R.M.;
 RT "Identification of a novel nuclear factor G1G2, as an ml-acetylcholine
 receptor-induced gene."
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF205437; AAG35663.1; -
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; pkinase; 2.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR ATP-binding; Receptor; Transferase.
 KW SEQUENCE 372 AA; 41008 MW; 5F54E50924B1365B CRC64;
 SQ

Query Match 38.9%; Score 735.5; DB 4; Length 372;
 Best Local Similarity 46.2%; Pred. No. 1.4e-54;
 Matches 160; Conservative 48; Mismatches 127; Indels 11; Gaps 6;

QY 2 RATPLAAPA-GSLSRKKRLLEDD--NLDTERPVQKRAASGP---QRLPFCLLPSPPA 55
 DB 17 RGPALLFPATKRCVPKRLDADDAVAAKCPRLSECCSPDYLSPPGSPC-SPOPPPA 75
 QY 56 PDRATAVATA--SRLGPFVLLPEPEGGRAYQALHCPGTETXCKVYPQALAVLEPYA 112
 DB 76 PGAGGSGSAPSPRSIADYLLPLAEREHVSALCIHGRRLRCVFPPIKHQDKIRPYI 135
 QY 113 RLPRKHVARPTEVLATGQLVAFTRTHGDMHSLVTRHRIPEPEAAVLFROMATALAH 172
 DB 136 QLPSSHSNITGIVEVILGETKAYVFPFKSGDMHYSVRSKRLREBEAARLFQIYSAVAH 195
 QY 173 CHQHGLVLRDLKLCGFVADDERKTLVLENDSCVLTPGPDLSLMDKACPAVYGPETLS 232
 DB 196 CHQSAIVIGDLKRFVFTSTERTQLRLESLEDTIMKGEDDALSCKGCPAYVSPETLN 255
 QY 233 SRASYSGRADAVMSLGVALLFTMLAGHYPFODESEPVLLFGKIRGAYALPAGISAPARCLV 292
 DB 256 TTGTYSGRADAVMSLGVALLFTMLAGHYPFODESEPVLLFGKIRGAYALPAGISAPARCLV 315
 QY 293 RCLLRPEPAERLTATGILLHFWLRQDPMPLAFTSRSHLMEAAQVVD 338
 DB 316 RSLLRPEPSERLTAPEILLHPFESVLEP-GYIDSEIGTSQIYVE 360

RESULT 10

Q91W04 PRELIMINARY; PRT; 372 AA.
 AC Q91W04;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Similar to phosphoprotein regulated by mitogenic pathways.
 GN TRB1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast tumor;
 RA Strauberg R.;
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Skin;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).

DR EMBL; EC006800; AA006800.1; -.
 DR EMBL; AK028626; BAC26038.1; -.
 DR MGI; MGI:2443397; Trb1.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; kinase; 2.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR ATP-binding; Transferase.
 KW SEQUENCE 372 AA; 41281 MW; AD29B84E640B462 CRC64;

Query Match 38.9%; Score 735.5; DB 11; Length 372;
 Best Local Similarity 46.2%; Pred. No. 1.4e-54;
 Matches 163; Conservative 46; Mismatches 119; Indels 25; Gaps 6;

QY 2 RATPLAAGSLSRKKRLDNDLTERPVQKARSGQPRLPCLLP---LSPPTAP-- 56
 DB 17 RGPGLLPFAAGTAPKRL-----LPTD--DAGAAVAAKCPRLSECSPPDYLSPPGSPCS 68
 QY 57 -----DATAVATASRLGP-----YVLLPEBEGRAVQALHCPGTGYTCVYVQVAL 105
 DB 69 PGPSTGTGSGCVSSPSPRIADYLLPLAERHVSALCIHGRERCKEPIKHYQ 128
 QY 106 AVLEPYARLPKHVARPTEVLAGTQLLYAFTTRHGMHSLVTRHRIPEEAAVLRQ 165
 DB 129 DKIRFYIQLPSSHSNTGIVGLVGLGSKAVVFEEKFGDMHSYVRSKRLREEAARLFQ 188
 QY 166 MATALAHCHQGLVLRDILCRFVADREKKLVLENLSDCVLTGPPDLSMDKACPAY 225
 DB 189 IVSAVAHCHQSAIVGDLKLRKFVSTERTQRLGSLIEDTHIIKGEDDALSDKHGCPAY 248
 QY 226 VGPETLSRASYSGKAADVWSLGVALLFTMLAGHYFODSEPVLLRGKIRGAYALPAGLS 285
 DB 249 VSPETLNTGTYSGKAADVWSLGVALLFTMLAGHYFODSEPVLLRGKIRGAYALPAGLS 308
 QY 286 APARCLVCLRLRREPARELTATGILHPMLRQDPMPPLATRSHLWEAAVVD 338
 DB 309 PPARCLIRSLRREPSERLTAQILLHPMFEVLEP-GYVDSIEIGTSQDIVE 360

RESULT 11
 Q8K4K4 PRELIMINARY; PRT; 372 AA.
 AC Q8K4K4;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, last annotation update)
 DE TRB-1.
 GN TRB1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kiss-Toch E., Dempsey C., Jozsa V., Gaunt J., Oxley K.M.,
 RA Bagstaff S.M., Wyllie D.H., Harre M., O'Neill L.A.D., Qvarnstrom E.E.,
 RA Power S.K.;
 RT "Mammalian homologs of Drosophila tritubles (trtb) control mitogen
 RT activated protein kinase signaling";
 RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF358866; AAM45478.1; -.
 DR MGI; MGI:2443397; Trb1.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00069; kinase; 2.

DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC_1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 KW SEQUENCE 372 AA; 41282 MW; 3A3DB82B46CD907F CRC64;

Query Match 38.2%; Score 722.5; DB 11; Length 372;
 Best Local Similarity 45.6%; Pred. No. 1.9e-53;
 Matches 161; Conservative 46; Mismatches 121; Indels 25; Gaps 6;

QY 2 RATPLAAGSLSRKKRLDNDLTERPVQKARSGQPRLPCLLP---LSPPTAP-- 56
 DB 17 RGPGLLPFAAGTAPKRL-----LPTD--DAGAAVAAKCPRLSECSPPDYLSPPGSPCS 68
 QY 57 -----DATAVATASRLGP-----YVLLPEBEGRAVQALHCPGTGYTCVYVQVAL 105
 DB 69 PGPSTGTGSGCVSSPSPRIADYLLPLAERHVSALCIHGRERCKEPIKHYQ 128
 QY 106 AVLEPYARLPKHVARPTEVLAGTQLLYAFTTRHGMHSLVTRHRIPEEAAVLRQ 165
 DB 129 DKIRFYIQLPSSHSNTGIVGLVGLGSKAVVFEEKFGDMHSYVRSKRLREEAARLFQ 188
 QY 166 MATALAHCHQGLVLRDILCRFVADREKKLVLENLSDCVLTGPPDLSMDKACPAY 225
 DB 189 IVSAVAHCHQSAIVGDLKLRKFVSTERTQRLGSLIEDTHIIKGEDDALSDKHGCPAY 248
 QY 226 VGPETLSRASYSGKAADVWSLGVALLFTMLAGHYFODSEPVLLRGKIRGAYALPAGLS 285
 DB 249 VSPETLNTGTYSGKAADVWSLGVALLFTMLAGHYFODSEPVLLRGKIRGAYALPAGLS 308
 QY 286 APARCLVCLRLRREPARELTATGILHPMLRQDPMPPLATRSHLWEAAVVD 338
 DB 309 PPARCLIRSLRREPSERLTAQILLHPMFEVLEP-GYVDSIEIGTSQDIVE 360

RESULT 12
 Q7ZZZ7 PRELIMINARY; PRT; 344 AA.
 AC Q7ZZZ7;
 DT 01-UN-2003 (TREMBLrel. 24, Created)
 DT 01-UN-2003 (TREMBLrel. 24, last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, last annotation update)
 DE TRB-2 kinase.
 GN Xenopus laevis (African clawed frog).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shan Y.X., Yu L.;
 RT "Cloning and characterization of the frog TRB-2 gene";
 RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AY254200; AAP13074.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0006473; F:protein tyrosine kinase activity; IEA.
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00069; kinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC_1.
 DR SMART; SM00219; TYKC; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 KW SEQUENCE 344 AA; 38867 MW; 900F4448EFAA468E CRC64;

Query Match 37.9%; Score 717; DB 13; Length 344;
 Best Local Similarity 46.7%; Pred. No. 4.9e-53;
 Matches 162; Conservative 45; Mismatches 122; Indels 18; Gaps 7;

QY 2 RANPLAPAGSLRKRRLDNLPTFRPYOKRARSQPRLPCLPLSPPTAPRATA 61
 Db 5 RSNPILITTYGRNRNKNQDPFE-LSSIRSTE-----PQSRSPNIGSSSPETPNSSHC 57
 QY 62 VATAASLGPVLLPEPEGGAVALHCPGTETVCKVYPV--OEALAVLEPYARLPK 118
 Db 58 V--SCIGKYLLEPLEGNHVPFALHLSGEFLCKVPFDICQETLA---PFCPLPIH 111
 QY 119 HVARPTEVLAGTOLYAFPTRHGDHSLVTRHRIPEEDAVLFFOMATLALHCHQGL 178
 Db 112 NINQIAEILIGELKAVFPERSHGDMHSFVCTCKEBAARLFQIVSAVAHCHDXGV 171
 QY 179 VLRLDLKCRFVFADEPRKKVLLENLEDSCVLTGPDLSLMDKACPPAYGPELSSRAYS 238
 Db 172 VLRLDLKRFVENDXRTYVKLSLSDAYVLSGSDLSDKGCPAYSPPELINTNGSYS 231
 QY 239 GKADVSLGVALFTMLAGHYPPDSEPVLLFGKIRGAYVALPAGLSAPAR-CLVYCLLR 297
 Db 233 GKADVMSLGVALFTMLVGRYFPHDIEPSSLFISKIRGQFNIPKILSPGQNVLYVAFRR 291
 QY 298 REPAPRLTATGILLHPWLRQDMPLAPITRSHLMEAAQVVPDGLGDE 344
 Db 292 REPSERLTSQETILDHPWFSTDPNALNSGCAKEVSDQVLPD-VNMDE 337

RESULT 13

Q8R2V8 PRELIMINARY; PRT; 257 AA.
 AC Q8R2V8;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, last annotation update)
 DE Hypothetical protein (Fragment).
 GN AW19517.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC027159; AAH27159.1; -
 DR MGD; MGI:2145021; AW19517
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0004674; F:protein kinase activity; IEA.
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; Pkinase; 2_
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KM Hypothetical protein; ATP-binding; Transferase.
 FT NON TER
 SQ SEQUENCE 257 AA; 29232 MW; 84BEACDD476F51 CRC64;

Query Match 34.7%; Score 657; DB 11; Length 257;
 Best Local Similarity 52.7%; Pred. No. 4.6e-48;

Matches 135; Conservative 33; Mismatches 68; Indels 20; Gaps 3;

QY 96 CKVYPV--OEALVLEPYARLPKHYARPTVLAGTOLYAFPTRHGDHSLVTRRH 152
 Db 3 CKVEISGYQSLA---PCFCLSAHSNNTQTEILLGTAKYVFFERSYGMHSFVRCK 59
 QY 153 RIPEEAAVLFROMATLALHCHQGLVRLDLKCRFVFADEPRKKVLLENLEDSCVLTGP 212
 Db 60 KLREEEARLFYQJASAVAHCHDGIVRLDLKRFKDEKRRVLSLSDAYVLRGD 119
 QY 213 DDSTLMDGACGAYVGPPELSSRAYSCKADVMSLGVALFTMLAGHYPPDSEPVLLFGK 272
 Db 120 DDSLSDKGGCPAYVSPPELINTNGSYSCKADVMSLGVALFTMLVGRYFPHDIEPSSIFS 179
 QY 273 IRRGAYVALPAGLSAPARCLVCLLRPEAPRLTATGILLHPWLRQDMPLAPITRSHLME 332

Db 180 IRRGQFNIPETLSPRAKCLIRSLRPESERLTSQETILDHPWFSTD----- 225
 QY 333 AQVVPDGLGDEAREE 348
 Db 226 FSVNSNGFKAACDQ 241

RESULT 14

O15180 PRELIMINARY; PRT; 224 AA.
 AC O15180;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JUN-1998 (TREMBLrel. 05, last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, last annotation update)
 DE Phosphoprotein (Fragment).
 GN C8FW.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Thyroid;
 RA Wilkain F.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE=98000262; PubMed=9342215;
 RA Wilkin F., Suarez-Huerta N., Robaye B., Peetermans J., Libert F.,
 RA Dumont J.E., Maenhaut C.;
 RT "Characterization of a phosphoprotein whose mRNA is regulated by the
 RT mitogenic pathways in dog thyroid cells.";
 RL Eur. J. Biochem. 248:660-669(1997).
 DR EMBL; AJ000480; CAA04119.1; -
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0004672; F:protein kinase activity; TAS.
 DR GO; GO:0008283; P:cell proliferation; TAS.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; Pkinase; 2_
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KM ATP-binding; Transferase.
 FT NON TER
 SQ SEQUENCE 224 AA; 25479 MW; 740CD0905F86499B CRC64;

Query Match 32.2%; Score 608.5; DB 4; Length 224;
 Best Local Similarity 56.2%; Pred. No. 5.4e-44;

Matches 118; Conservative 32; Mismatches 59; Indels 1; Gaps 1;

QY 129 GTOLLVAFPTRHGDHSLVTRRHRIPEEAAVLFROMATLALHCHQGLVRLDLKCRF 188
 Db 4 GETAYVFFEKSRFDMHSYASRKRLRBEERARLFKQIVSAVAHCHQSAIVLGLDKRF 63
 QY 189 VFADREKKVLLENLEDSCVLTGPDLSLMDKACPPAYGPELSSRAYSCKADVMSLG 248
 Db 64 VFSTEERTQRLSELLEDTHIMKGEDDALSCHGCPAYSPPELINTNGSYSCKADVMSLG 123
 QY 249 VALFTMLAGHYPPDSEPVLLFGKIRGAYVALPAGLSAPARCLVCLLRPEAPRLTATG 308
 Db 124 VMLYTLVGRYFPHDSDPSALFSKIRGQFCIPRHISPKACILRLRRESERLTAP 183
 QY 309 ILHPWLRQDMPLAPITRSHLMEAAQVVPD 338
 Db 184 ILHPWFESVLEP-GYIDSEIGSDQIVPE 212

RESULT 15

O9V321 PRELIMINARY; PRT; 484 AA.
 AC O9V321;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 29, 2004, 04:38:36 ; Search time 831 Seconds

(without alignments)
10817.313 Million cell updates/sec

Title: US-10-070-337-16
Perfect score: 2116

Sequence: 1 ggaagcgagcttcgcgcgcgt.....tatgaataaaaaaaaaa 2116

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N Geneseg 29Jan04:*

1: Geneseg1980s:*\n2: Geneseg1990s:*\n3: Geneseg2000s:*\n4: Geneseg2001as:*\n5: Geneseg2001bs:*\n6: Geneseg2002s:*\n7: Geneseg2003as:*\n8: Geneseg2003bs:*\n9: Geneseg2003cs:*\n10: Geneseg2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2116	100.0	2116	AAI30480	Human pro
2	2095.4	99.0	2257	ADCS9338	DNA seque
3	2039.4	96.4	2092	AAI58064	Human pol
4	2039.4	96.3	2059	AAI59850	Human pol
5	1762.4	83.3	2048	AAI7866	Human can
6	1072.2	50.7	1077	AAI506709	Human can
7	1069.8	49.8	1085	AAI261155	Polynucle
8	1063.8	49.8	1083	AAI86479	Human tri
9	803.8	38.0	972	AAH76218	Human kin
10	716.6	33.9	1076	ABH39747	Human NS
11	716.6	33.9	1076	ABH39747	Human NS
12	621.6	29.4	2004	ADB58995	Human NS
13	621.6	29.4	2004	ADB58995	Human NS
14	621.6	29.4	2004	ADB58995	Human NS
15	562.4	25.2	996	AAI591231	DNA seque
16	533	25.2	996	AAI591231	DNA seque
17	505.2	23.9	2559	ADB31419	Human col
18	505.2	23.9	2559	ADB31419	Human col
19	426	20.1	426	AAQ89817	Biclutam
20	369.4	17.5	426	ABH87304	DNA encod
21	354.2	16.7	366	AAI78131	Human ova
22	315.4	14.9	353	AAI37461	Human can
23	297.6	14.1	534	AAI62004	Colon ade
				AAI66600	DNA encod

24	269.4	12.7	3319	ABN86478	Human tri
25	267.8	12.7	3328	ADD29847	Human tum
26	258.4	12.2	1909	AAI45035	CDNA enco
27	251.4	11.9	4221	ADD14683	Human src
28	249.8	11.8	3280	ABX63269	Human src
29	243.8	11.5	3324	AAI594913	Human DNA
30	240	11.3	690	ABT09024	Human DNA
31	233.4	11.0	942	ABX74428	Phase-1 R
32	230.8	10.9	675	ABK83550	Human CDN
33	229.2	10.8	818	ACC45125	Human CDN
34	219	10.3	621	ACC45126	Human CDN
35	205	9.7	207	AAI03717	Human CBF
36	204.4	9.7	1943	AAI45223	Human sec
37	165	7.8	498	ABK63084	CDNA enco
38	163.6	7.7	7282	AAI86318	Rat seque
39	158	7.5	562	AAI41844	Human imm
40	158	7.5	562	AAI41844	Human fce
41	158	7.5	562	AAI41844	Prode #10
42	158	7.5	562	AAI41844	Human Don
43	158	7.5	562	AAI41844	Human bra
44	158	7.5	562	ABK10231	Human liv
45	158	7.5	562	ABK10267	Human gen
				AAI81332	Human imm

ALIGNMENTS

RESULT 1	AAI30480	standard; cDNA; 2116 BP.
ID	AAI30480	
XX	AAI30480	
AC	AAI30480	
XX	AAI30480	
DT	29-MAY-2001	(first entry)
XX	AAI30480	
DE	Human protein phosphatase and kinase protein-5 cDNA 1271505CBL1.	
XX	AAI30480	
KW	Protein phosphatase and kinase protein; PPKP-5; human;	
KW	gastrointestinal disorder; immune system disorder; neurological disorder;	
KW	cell proliferative disorder; cancer; diagnosis; therapy; ss.	
XX	AAI30480	
OS	Homo sapiens.	
XX	AAI30480	
FX	Key	Location/Qualifiers
FT	CDS	161..1237
FT	misc_feature	326..385
FT	FT	/*tag= b
FT	FT	/note= "unique fragment"
XX	AAI30480	
PN	W0200120004-A2.	
XX	AAI30480	
PD	22-MAR-2001.	
XX	AAI30480	
PP	14-SEP-2000; 2000WO-US025515.	
XX	AAI30480	
PR	15-SEP-1999; 99US-0154141P.	
XX	AAI30480	
PA	(INCY-) INCYTE GENOMICS INC.	
XX	AAI30480	
PI	Yue H, Tang YT, Bandman O, Hallman JL, Baughn MR, Azimzai Y;	
PI	Lu DM;	
XX	AAI30480	
DR	WPI; 2001-244811/25.	
DR	P-RSDB; AAB20326.	
XX	AAI30480	
PT	Novel human protein phosphatase and kinase proteins for diagnosis,	
PT	treatment and prevention of gastrointestinal, immune system, neurological	
PT	and cell proliferative disorders.	
XX	AAI30480	
PS	Claim 5; Page 98; 103pp; English.	
XX	AAI30480	
CC	The present sequence is that of cDNA encoding novel human protein	

QY 1861 GTTCTGTAACCTGAAGACTGTTCTGGAATGAGGGTCCAGGCTCTGCAACCTGAGGCT 1920
 Db 1861 GTTCTGTAACCTGAAGACTGTTCTGGAATGAGGGTCCAGGCTCTGCAACCTGAGGCT 1920
 QY 1921 TCTGACCTGAGCACCAGGTTGAGGAGCAGATTTAGGAGGGCTCTGCTGGGACCT 1980
 Db 1921 TCTGACCTGAGCACCAGGTTGAGGAGCAGATTTAGGAGGGCTCTGCTGGGACCT 1980
 QY 1981 GGAAGTCCAGGTTGAGGAGCTTTCTGGGAGCAGCTTGGGGTCCACATCCAGGTCATAC 2040
 Db 1981 GGAAGTCCAGGTTGAGGAGCTTTCTGGGAGCAGCTTGGGGTCCACATCCAGGTCATAC 2040
 QY 2041 TCTAGGTTTGAATCCATGATGATGATTTACTGTCCTAATAAAGAGATATG 2100
 Db 2041 TCTAGGTTTGAATCCATGATGATGATTTACTGTCCTAATAAAGAGATATG 2100
 QY 2101 AATTAATAAAAAAAAAA 2116
 Db 2101 AATTAATAAAAAAAAAA 2116

RESULT 2

ADCS9338
 ID ADCS9338 standard; DNA; 2257 BP.

AC ADCS9338;

DT 18-DEC-2003 (first entry)

DE DNA sequence encoding human cell death inhibitory protein.

XX Human; cell-death; cardiact; neuroprotective; anti-HIV; antiinflammatory;
 XX cerbrioprotective; hepatotropic; cytostatic; immunosuppressive;
 XX antirheumatic; ophthalmological; nootropic; antiparkinsonian;
 XX anticonvulsant; hypotensive; antiarteriosclerotic; haemostatic;
 XX antialcoholic; virucide; HIV; cardiac disease; immunological disease;
 XX neurodegenerative disease; ischaemic damage;
 XX central nervous system disorder; hepatic disorder;
 XX acute pancreatic inflammation; cancer; AIDS; autoimmune disease;
 XX rheumatoid; Crohn's disease; glaucoma; Alzheimer's disease;
 XX Parkinson's disease; Huntington's disease; hypertension;
 XX arteriosclerosis; reperfusion damage; myocardial infarction;
 XX cerebral trauma; cerebral infarction; cerebral haemorrhage; hepatitis;
 XX alcoholic hepatitis; cerebral ischemia; ss; gene.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 295..1371

FT /*tag= a

PN /product= "Human cell death inhibitory protein"

JP2003063986-A.

PD 05-MAR-2003.

PF 27-AUG-2001; 2001JP-00255811.

PR 27-AUG-2001; 2001JP-00255811.

PA (SUMU) SUMITOMO SEIYAKU KK.

XX WPI; 2003-601360/57.

DR P-PSDB; ADCS9337.

PT Controlling cell-death by administering positive or negative regulator of

PT cell-death inhibitory-factor, for treating cancer; AIDS; autoimmune

PT diseases; Crohn's diseases; glaucoma; Alzheimer disease.

PS Claim 1; SEQ ID NO 4; 35pp; Japanese.

XX This invention relates to a novel method for controlling cell-death

CC comprising administering a positive or negative regulator of cell-death
 CC inhibitory-factor. The invention also comprises a method for screening
 CC modulators of cell-death, by contacting cells expressing cell-death
 CC inhibitory factor with a candidate compound, monitoring level of
 CC expression of cell-death inhibitory factor, evaluating cell-death
 CC modulation ability of the compound based on change in the level of
 CC expression of the factor and selecting compounds having cell-death
 CC modulation ability. The cell death regulator of the invention may have
 CC cardiact, neuroprotective, anti-HIV, antiinflammatory, cerebroprotective,
 CC hepatotropic, cytostatic, immunosuppressive, antirheumatic,
 CC ophthalmological, nootropic, antiparkinsonia, anticonvulsant,
 CC hypotensive, antiarteriosclerotic, haemostatic, antialcoholic and
 CC virucide activities. The method of the invention is useful for treating
 CC HIV, cardiac diseases, immunological diseases, neurodegenerative disease,
 CC ischaemic damage and congestion, disorder of central nervous system,
 CC hepatic disorder, acute pancreatic inflammation, and cancer, AIDS,
 CC autoimmune diseases, rheumatism, Crohn's disease, glaucoma, Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, hypertension,
 CC arteriosclerosis, reperfusion damage, myocardial infarction, cerebral
 CC trauma, cerebral infarction, cerebral haemorrhage, hepatitis, alcoholic
 CC hepatitis, and cerebral ischemia. The present sequence represents the DNA
 CC sequence encoding the human cell death inhibitory protein of the
 CC invention.

SQ Sequence 2257 BP; 435 A; 705 C; 659 G; 458 T; 0 U; 0 Other;

Query Match 99.0%; Score 2095.4; DB 9; Length 2257;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2110; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 GAGAGGAGCTCCGCGCGGCTCGCTGAGACCCGAGGAGCTGAGAGCTGGGCTGGGA 60
 Db 134 GAGAGGAGCTCCGCGCGGCTCGCTGAGACCCGAGGAGCTGAGAGCTGGGCTGGGA 193
 QY 61 TCCGAGGTTGGGAG 119
 Db 194 TCCGAGGTTGGGAG 253
 QY 120 GCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 179
 Db 254 GCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 313
 QY 180 CTGCTCTCGGAGGTTCTCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 239
 Db 314 CTGCTCTCGGAGGTTCTCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 373
 QY 240 CCGAGGCTCCGCTCCAG 299
 Db 374 CCGAGGCTCCGCTCCAG 433
 QY 300 TGTGACCTGAGGCCACTACTGCTCCAGATCGTGAACACTGTGTGAGCACTGCTGCC 359
 Db 434 TGTGACCTGAGGCCACTACTGCTCCAGATCGTGAACACTGTGTGAGCACTGCTGCC 493
 QY 360 GTCTTGGGCTTATGTCCTCTGTGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
 Db 494 GTCTTGGGCTTATGTCCTCTGTGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 553
 QY 420 ACTGACCTTACAGGAG 479
 Db 554 ACTGACCTTACAGGAG 613
 QY 480 TGTGAGAGCCCTTACAGGAG 539
 Db 614 TGTGAGAGCCCTTACAGGAG 673
 QY 540 TGGCTGGTACCCAGCTCTCTAGAGCCCTTTTCACTCGAGAGCCATGGAGAGAGAGAGAG 599
 Db 674 TGGCTGGTACCCAGCTCTCTAGAGCCCTTTTCACTCGAGAGCCATGGAGAGAGAGAGAG 733
 QY 600 TGGTGGAGAGGAGGAG 659
 Db 734 TGGTGGAGAGGAGGAG 793

QY 660 CCACCGCCCTGGCGCAGTGTCAACGACAGCTTGTCTCTGCTGATCTCAAGCTGTGTG 719
 Db 794 CCACCGCCCTGGCGCAGTGTCAACGACAGCTTGTCTCTGCTGATCTCAAGCTGTGTG 853
 QY 720 GCTTGTCTGCTGCTGACCGCTGAGAGAGAACTGTGTGTGAGAACTGAGAACTGCT 779
 Db 854 GCTTGTCTGCTGCTGACCGCTGAGAGAGAACTGTGTGTGAGAACTGAGAACTGCT 913
 QY 780 GCGTGTGA CTGGGCGCAGATGATTCCTGTGTGGAACAAGACGCTGTCCAGCTTACGTGG 839
 Db 914 GCGTGTGA CTGGGCGCAGATGATTCCTGTGTGGAACAAGACGCTGTCCAGCTTACGTGG 973
 QY 840 GACCTGAGATCTACAGCTACAGGGGCTCTACTCTGGGGACAGGACGCAATGTCTGAGGC 899
 Db 974 GACCTGAGATCTACAGCTACAGGGGCTCTACTCTGGGGACAGGACGCAATGTCTGAGGC 1033
 QY 900 TGGGCGTGGCGCTCTTCAACATGTGTGCGGCGCACTACCCCTTCCAGACTGAGAGCTG 959
 Db 1034 TGGGCGTGGCGCTCTTCAACATGTGTGCGGCGCACTACCCCTTCCAGAGCTGAGAGCTG 1093
 QY 960 TCTGTCTCTTGGGCAAGATCCGCGCGGCGCTACGCTTGTGCTGCAAGCTTGTGGGCC 1019
 Db 1094 TCTGTCTCTTGGGCAAGATCCGCGCGGCGCTACGCTTGTGCTGCAAGCTTGTGGGCC 1153
 QY 1020 CTGGCCGCTGT 1079
 Db 1154 CTGGCCGCTGT 1213
 QY 1080 CAGGCAATCTCTCTGACACCCCTGGCTGTGACAGGACCCGATGCTTGTGCTTGTGCTTGT 1139
 Db 1214 CAGGCAATCTCTCTGACACCCCTGGCTGTGACAGGACCCGATGCTTGTGCTTGTGCTTGT 1273
 QY 1140 CCCATCTCTGGGAGGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1199
 Db 1274 CCCATCTCTGGGAGGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1333
 QY 1200 AAGAGAGAGAGACAGAGAAAGTGTCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1259
 Db 1334 AAGAGAGAGAGACAGAGAAAGTGTCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1393
 QY 1260 CTGCAACAGT 1319
 Db 1394 CTGCAACAGT 1453
 QY 1320 CCAAACTTTCAGT 1379
 Db 1454 CCAAACTTTCAGT 1513
 QY 1380 TACACATCTGT 1439
 Db 1514 TACACATCTGT 1573
 QY 1440 GCGCTGTCTCGGT 1499
 Db 1574 GCGCTGTCTCGGT 1633
 QY 1500 AGAGATGACAACTGGGATCTTGT 1559
 Db 1634 AGAGATGACAACTGGGATCTTGT 1693
 QY 1560 CTACACATCTGT 1619
 Db 1694 CTACACATCTGT 1753
 QY 1620 CTGTCCAGAGACAACTCCCTTGT 1679
 Db 1754 CTGTCCAGAGACAACTCCCTTGT 1813
 QY 1680 AGAGAAAGGAGATCTCTGT 1739
 Db 1814 AGAGAAAGGAGATCTCTGT 1873

QY 1740 CAAGCCAGCTCACTCTGGGAAGT 1799
 Db 1874 CAAGCCAGCTCACTCTGGGAAGT 1933
 QY 1800 TCTCTTCCAGGCTTAACTCAAGACTGT 1859
 Db 1934 TCTCTTCCAGGCTTAACTCAAGACTGT 1993
 QY 1860 AGTCTGT 1919
 Db 1994 AGTCTGT 2053
 QY 1920 TTCTGACCTGAGCACAAGGTTGT 1979
 Db 2054 TTCTGACCTGAGCACAAGGTTGT 2113
 QY 1980 TGGAAAGTCCAGGTTGT 2039
 Db 2114 TGGAAAGTCCAGGTTGT 2173
 QY 2040 CTCTAGGTTTGT 2099
 Db 2174 CTCTAGGTTTGT 2233
 QY 2100 GAAATTAATAAAAAAAAAA 2116
 Db 2234 GAAATTAATAAAAAAAAAA 2250

RESULT 3
 ID AAI58064
 XX AAI58064 standard; cDNA; 2092 BP.
 AC AAI58064;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 267.
 XX
 KW Human; neotropic; immunosuppressant; cytosolic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00486725.
 PR 25-APR-2000; 2000US-00523117.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSEQ-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D, Zhao J, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Q, Zhou P, Goodrich R, Dmanac R;
 PI
 XX
 DR WPI; 2001-442253/47.
 DR P-PDB; AAM38908.
 XX

QY 1838 TAAGATCAAACTATGAGGCTAGTCTTCTTCTAACTCAAGCTGTTCTGAAATGAGGCT 1897
DB 1816 TAAGATCAAACTATGAGGCTAGTCTTCTTCTAACTCAAGCTGTTCTGAAATGAGGCT 1875
QY 1898 CAGAGCTGTCAACATGAGGCTTCTGACCTGAGACCAAGGTTGAGGACAGATTAAG 1957
DB 1876 CAGAGCTGTCAACATGAGGCTTCTGACCTGAGACCAAGGTTGAGGACAGATTAAG 1935
QY 1958 CAGAGCTGTCTGTGGCCACCTGGAAGTCCAGAGTGGACCTCTTCTGGGACACTTGG 2017
DB 1936 CAGAGCTGTCTGTGGCCACCTGGAAGTCCAGAGTGGACCTCTTCTGGGACACTTGG 1995
QY 2018 GGTCCCAATCCAGAGTCCATATCTAGGTTTGGATACATGATATGATTTACT 2077
DB 1996 GGTCCCAATCCAGAGTCCATATCTAGGTTTGGATACATGATATGATTTACT 2055
QY 2078 GGTCCCAATCCAGAGTCCATATCTAGGTTTGGATACATGATATGATTTACT 2114
DB 2056 GGTCCCAATCCAGAGTCCATATCTAGGTTTGGATACATGATATGATTTACT 2092

RESULT 4
AA159850
ID AA159850 standard; cDNA; 2059 BP.
AC AA159850;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 3839.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW Leukemia; ss.
XX
OS Homo sapiens.
XX
PN WO20015312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-0055317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Q;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR F-P-SDB; AAM40694.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 3839; 10078bp; English.
CC The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localized neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 2059 BP; 400 A; 645 C; 583 G; 431 T; 0 U; 0 Other;
XX
Query Match 96.3%; Score 2038.6; DB 4; Length 2059;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2052; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 59 GATCCCGAGCTCCGAGAGAGGCA-CGGGCGGCGCCACCTGCTGCTGAGAGGCTCT 117
DB 3 GATCCCGAGCTCCGAGAGAGGCA-CGGGCGGCGCCACCTGCTGCTGAGAGGCTCT 62
QY 118 GAGCCCCGCGCGCCGCGGCGCAAGCGGGAAGAGGCGAGATGAGAGCCACCTCT 177
DB 63 GAGCCCCGCGCGCGCGGCGCCAGCGGGAAGAGGCGAGATGAGAGCCACCTCT 122
QY 178 GAGCTCTCTGCGGCTTCTCTGTCAGAGAAAGCGGTTGAGATTGATGAACTTAGA 237
DB 123 GAGCTCTCTGCGGCTTCTCTGTCAGAGAAAGCGGTTGAGATTGATGAACTTAGA 182
QY 238 TACCGAGAGTCCCGCCGCAAGAAAGAGTGGGCGCCAGACCTGCCCCCTG 297
DB 183 TACCGAGAGTCCCGCCGCAAGAAAGAGTGGGCGCCAGACCTGCCCCCTG 242
QY 298 CCGTGTGCGCCGAGACCACTACTGCTCCAGATGTCGCACTGTGGCCACTGCTC 357
DB 243 CCGTGTGCGCCGAGACCACTACTGCTCCAGATGTCGCACTGTGGCCACTGCTC 302
QY 358 CCGTGTGCGCCGCTATGTCCTCTGAGAGCCGAGAGGCGGCGCTTACCAAGCCCT 417
DB 303 CCGTGTGCGCCGCTATGTCCTCTGAGAGCCGAGAGGCGGCGCTTACCAAGCCCT 362
QY 418 GCATGCGCTTACAGGACCTGATATACCTGCAAGGTTGATCCCGTCAGAGAGCCCTGCG 477
DB 363 GCATGCGCTTACAGGACCTGATATACCTGCAAGGTTGATCCCGTCAGAGAGCCCTGCG 422
QY 478 CGTGTGAGAGCCCTAAGCGCGCTGCGCCCGCAAGAGATGTGGCTGCGCCACTGAGGT 537
DB 423 CGTGTGAGAGCCCTAAGCGCGCTGCGCCCGCAAGAGATGTGGCTGCGCCACTGAGGT 482
QY 538 CCGTGTGAGTACCAAGCTCTCTAAGCTTTTCACTCGGACCCATGAGGACATGTAGAG 597
DB 483 CCGTGTGAGTACCAAGCTCTCTAAGCTTTTCACTCGGACCCATGAGGACATGTAGAG 542
QY 598 CCGTGTGAGAGCGCGCAAGCTATCCCTGAGCTGAGCTGCGCTCTTCCGACAGAT 657
DB 543 CCGTGTGAGAGCGCGCAAGCTATCCCTGAGCTGAGCTGCGCTCTTCCGACAGAT 602
QY 658 GGCACACCGCCTGAGCGCACTGTACAGAGAGGAGTGTCTCGGTGATCTCAAGCTGTG 717
DB 603 GGCACACCGCCTGAGCGCACTGTACAGAGAGGAGTGTCTCGGTGATCTCAAGCTGTG 662
QY 718 TCGCTTTGCTTGTGCTGACCCGTGAGAGGAAAGCTGTGTCTGAGAACTGAGAGACTC 777
DB 663 TCGCTTTGCTTGTGCTGACCCGTGAGAGGAAAGCTGTGTCTGAGAACTGAGAGACTC 722
QY 778 CTGCGGTGAGCTGAGGCGAGATATCTCTGTGAGCAAGAGCGCTGCGACCTTACGT 837
DB 723 CTGCGGTGAGCTGAGGCGAGATATCTCTGTGAGCAAGAGCGCTGCGACCTTACGT 782
QY 838 GGCAGCTGAGATCTCAGCTCAGGCGCTCATATCTGCGGCAAGGCAAGCGATGTCTGAG 897

Db	783	GGACACTGGAGTAACTCAGCTCAACGGGCTCTAATCTCGGCAAGCAGCGAGTGTCTGGAG	842
OY	898	CCGAGGACGTGGGCGCTTTACCATGGCTGGCGGGCCATACCCCTTCAGAGATTCGGAGCGC	957
Db	843	CTGGGCGCTGGCGCTTTACCAATGTGGCGGGCCATACCCCTTCAGAGATTCGGAGCGC	902
OY	958	TGTCTGTCTTTTGGCAAGATCGCGCGGGGCTTACGCTTTCCTGCAAGCGCTCTGGAGC	1011
Db	903	TGTCTGTCTTTTGGCAAGATCGCGCGGGGCTTACGCTTTCCTGCAAGCGCTCTGGAGC	962
OY	1018	CCCTGCGCGCTGTCTGGTTCGTGCTCTTGTGTGGGAGCCACTGAACTGAGCTCTGAGC	1077
Db	963	CCCTGCGCGCTGTCTGGTTCGTGCTCTTGTGTGGGAGCCACTGAACTGAGCTCTGAGC	1022
OY	1078	CACAGGACATCCCTCGGCAACCCCTGGCTGGACAGAACCCGATAGGCCCTTAAGGCCCAACCG	1133
Db	1023	CACAGGACATCCCTCGGCAACCCCTGGCTGGACAGAACCCGATAGGCCCTTAAGGCCCAACCG	1082
OY	1138	ATCCCATCTTGGGAGGCTGCCGAGGTGTCCTGATGATGATGGGAGCTGGAGGCTGACGAGCCGAC	1197
Db	1083	ATCCCATCTTGGGAGGCTGCCGAGGTGTCCTGATGATGATGGGAGCTGGAGGCTGACGAGCCGAC	1142
OY	1198	GGAAAGAGAGGGGAGACAGAGAAAGTGGTCTGTAAATGGCTAAGAACCCCTTAACACCTC	1251
Db	1143	GGAAAGAGAGGGGAGACAGAGAAAGTGGTCTGTAAATGGCTAAGAACCCCTTAACACCTC	1202
OY	1258	AGCTGCCAACAGTGGATTGAGTTGGGGGTAGCTTCAAGCCCTTCTCGCTCTGAACTG	1317
Db	1203	AGCTGCCAACAGTGGATTGAGTTGGGGGTAGCTTCAAGCCCTTCTCGCTCTGAACTG	1262
OY	1318	AGCCAAACCTTCAGTGCCTTCCAGAAAGGAGAAAGCCCTGTGTGATGTGTCTG	1377
Db	1263	AGCCAAACCTTCAGTGCCTTCCAGAAAGGAGAAAGCCCTGTGTGATGTGTCTG	1322
OY	1378	TGTACACATCTGCTTTGTTCACACACATGCAAGTTCCTGTTGGGTGTATTCAGGTGCC	1437
Db	1323	TGTACACATCTGCTTTGTTCACACACATGCAAGTTCCTGTTGGGTGTATTCAGGTGCC	1382
OY	1438	AAGCCCTGTTCGGTGTGGGAGTACGACAGTGAAGCAAGAGAACATATTTCCCTGCTC	1497
Db	1383	AAGCCCTGTTCGGTGTGGGAGTACGACAGTGAAGCAAGAGAACATATTTCCCTGCTC	1442
OY	1498	ACAGAGATGACAAACTGGGCACTCTGAGCTGACAACTTTTGCATGACCTATAGTCACT	1557
Db	1443	ACAGAGATGACAAACTGGGCACTCTGAGCTGACAACTTTTGCATGACCTATAGTCACT	1502
OY	1558	GTCATCACTGGGTACACTTTGTATACCAAGTGTGGGCTCCACATGATGTGATCTCAGGAC	1617
Db	1503	GTCATCACTGGGTACACTTTGTATACCAAGTGTGGGCTCCACATGATGTGATCTCAGGAC	1562
OY	1618	CTGTGTCCAGAGACAAATCCCTTTTCAAAACAACAAGTGTGCTTTGTATCTTGAACCTTT	1677
Db	1563	CTGTGTCCAGAGACAAATCCCTTTTCAAAACAACAAGTGTGCTTTGTATCTTGAACCTTT	1622
OY	1678	TCAAGAGAAAGGAGGTATCCCTGTGCCAAAGGCTCCAGGCTCCCTCCGTGAATCCAGAA	1737
Db	1623	TCAAGAGAAAGGAGGTATCCCTGTGCCAAAGGCTCCAGGCTCCCTCCGTGAATCCAGAA	1682
OY	1738	CCCAAGCCAGCTCACTCTGGGAACTGTTCAGAGACCTCTGCTCTTGTGATTAAGAG	1797
Db	1683	CCCAAGCCAGCTCACTCTGGGAACTGTTCAGAGACCTCTGCTCTTGTGATTAAGAG	1742
OY	1798	ATTCTCCCTCCAGGCTTAAGCTGGGAAATTTGGGCGACAGATTAAGAAATTCGAATCTATAGG	1857
Db	1743	ATTCTCCCTCCAGGCTTAAGCTGGGAAATTTGGGCGACAGATTAAGAAATTCGAATCTATAGG	1802
OY	1858	CTAGTCTTGTCTTAACCAAGACTGTCTGGAAATGAAGGTCCAGGCTGTCAACATAGG	1917
Db	1803	CTAGTCTTGTCTTAACCAAGACTGTCTGGAAATGAAGGTCCAGGCTGTCAACATAGG	1862
OY	1918	GCTTCTAACCTGACCAACAGTTGAAGGACAGGATTAAGGAGGCTGTCTGTGGCCA	1977

Db	1863	GCTTCTGACCTGACACCAAGGTTGAG86GACAGATTAGAGAGGCTCTGTCCTGTGGCA	1922
QY	1978	CCTGGAAGTCCCGAGGTGGACACTTTCTGGGAGACCTTGGGGTCCACAAATCCAGGTCCA	2037
Db	1923	CCTGGAAGTCCCGAGGTGGACACTTTCTGGGAGACACTTGGGGTCCACAAATCCAGGTCCA	1982
QY	2038	TACCTCTAGGTTTGATATACATGATGATATGTTTACCTGTGGCTTAATAGAGAAAT	2097
Db	1983	TACCTCTAGGTTTGATATACATGATGATATGTTTACCTGTGGCTTAATAGAGAAAT	2042
QY	2098	ATGAAATPAAAAAAAAA 2114	
Db	2043	ATGAAATPAAAAAAAAA 2059	
RESULT 5			
AACT7866			
ID	AACT7866	standard; cDNA; 2048 BP.	
AC	AACT7866;		
XX	08-FEB-2001	(first entry)	
DT			
DE	Human cancer associated gene sequence SEQ ID NO:260.		
XX			
XX	Human; cancer associated gene; cancer antigen; detection; cancer;		
KW	diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;		
KW	antidiabetic; antisthmatic; antirheumatic; antiarthritic; antiviral;		
KW	antiinflammatory; antihypoid; antiallergic; antibacterial; cardiant;		
KW	dermatological; neuroprotective; thrombolytic; coagulant; noctropic;		
KW	vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;		
KW	immune disorder; haematopoietic cell disorder; autoimmune disorder;		
KW	allergic reaction; graft versus host disease; organ rejection;		
KW	haemostatic; thrombolytic; cardiovascular disorder; infection;		
KW	neurological disease; drug screening; ss.		
XX			
XX	Homo sapiens.		
OS			
XX	WO20005350-A1.		
PN			
PD	21-SEP-2000.		
XX			
PF	08-MAR-2000; 2000WO-US005882.		
XX			
PR	12-MAR-1999; 99US-0124270P.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI			
PI	Rosen CA, Ruben SM;		
XX	WPI; 2000-587533/55.		
DR	P-PSDB; AAB43657.		
XX			
XX	Nevel isolated nucleic acids comprising sequences encoding peptides		
XX	useful for treating or diagnosing e.g. cancer.		
PS			
XX	Claim 1, Page 825-826; 2352pp; English.		
CC	AACT7607 to AACT8448 encode the human cancer associated proteins given in		
CC	AAB43398 to AAB44239. The proteins can have activities based on the		
CC	tissues and cells the genes are expressed in. Example of activities		
CC	include: cytostatic; proliferative; vulneray; immunomodulator;		
CC	antidiabetic; antisthmatic; antirheumatic; antiarthritic;		
CC	antiinflammatory; antihypoid; antiallergic; antibacterial; antiviral;		
CC	dermatological; neuroprotective; cardiant; thrombolytic; coagulant;		
CC	noctropic; vasotropic; antipsoriatic and antiangiogenic. The		
CC	polynucleotides and polypeptides can be used for preventing, treating or		
CC	ameliorating medical conditions and diagnosing pathological conditions.		
CC	Polynucleotides, polypeptides, antibodies, agonists and antagonists from		
CC	the present invention may be used to treat immune disorders by activating		
CC	or inhibiting the proliferation, differentiation or mobilisation of		
CC	immune cells, to treat disorders of haematopoietic cells, autoimmune		
CC	disorders, allergic reactions, graft versus host disease and organ		


```

RESULT 6
AAS06709
ID AAS06709 standard; cDNA; 1077 BP.
XX
AC AAS06709;
XX
DT 12-SEP-2001 (first entry)
XX
DE Polynucleotide sequence encoding human protein kinase #9.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US032085.
XX
PR 24-NOV-1999; 99US-0167482P.
XX
PA (SUGB-) SUGEN INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clardy D;
XX
DR WPI; 2001-343950/36.
XX
P-PSDB; AAN03509.
XX
PT Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections.
XX
PS Example 1; Fig 1; 433bp; English.
XX
CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel
CC protein kinases have been identified as members of the tyrosine or
CC serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be used for
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC polypeptides may be used as antigens in the production of antibodies
CC against the protein kinases and in assays to identify modulators of
CC protein kinase expression and activity
XX
SQ Sequence 1077 BP; 174 A; 373 C; 325 G; 205 T; 0 U; 0 Other;
Query Match 50.7%; Score 1072.2; DB 4; Length 1077;
Best Local Similarity 99.7%; Pred. No. 2e-259;
Matches 1074; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 161 ATGGAGGCGACCCCTTGGCTGCTCTGGCGGGTTCCCTGTCAGAGAAAGCGGTGGAG 220
DB 1 ATGGAGGCGACCCCTTGGCTGCTCTGGCGGGTTCCCTGTCAGAGAAAGCGGTGGAG 60
QY 221 TTGGATGACAACTAGATACCGAGCGTCCCTCCAGAAAGAGCTCCAGAGTGGCCCGCAG 280
DB 61 TTGGATGACAACTAGATACCGAGCGTCCCTCCAGAAAGAGCTCCAGAGTGGCCCGCAG 120
QY 281 CCCAGACTGCCCCCTGCTGTGCTCCCTGAGCCCACTACTGCTCCAGATCGTGCAGCT 340

```

```

DB 121 CCCAGACTGCCCCCTGCTGTGCTCCCTGAGCCCACTACTGCTCCAGATCGTGCAGCT 180
QY 341 GCTGTGGCCACTGCTCCCTGCTGTGGGCTTATGTCCTCTGGAGCCGAGAGGGCCGG 400
DB 181 GCTGTGGCCACTGCTCCCTGCTGTGGGCTTATGTCCTCTGGAGCCGAGAGGGCCGG 240
QY 401 CGGGGCTTACAGAGCCCTGCTGCTTACAGGCACTGATATACCTGCAAGGTATACCC 460
DB 241 CGGGGCTTACAGAGCCCTGCTGCTTACAGGCACTGATATACCTGCAAGGTATACCC 300
QY 461 GTCCAGAGAAAGCCCTGCTGCTGAGACCTTACCGCGGCTGCCCGCCAGAGCATGTG 520
DB 301 GTCCAGAGAAAGCCCTGCTGCTGAGACCTTACCGCGGCTGCCCGCCAGAGCATGTG 360
QY 521 GCTCGAGCCCACTGAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 580
DB 361 GCTCGAGCCCACTGAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 581 CATGGGGAACATGACAGGCTGTGGGAAAGCGCCCAACCGTATCCCTGAGGCTGAGC 640
DB 421 CATGGGGAACATGACAGGCTGTGGGAAAGCGCCCAACCGTATCCCTGAGGCTGAGC 480
QY 641 GTGCTCTTCCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 700
DB 481 GTGCTCTTCCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 701 CGTATCTCAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 760
DB 541 CGTATCTCAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 761 GAGAACTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 820
DB 601 GAGAACTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 821 GCGTCCCAAGCTTACCTGAGGACCTGAGATCTGAGATCTGAGATCTGAGATCTGAG 880
DB 661 GCGTCCCAAGCTTACCTGAGGACCTGAGATCTGAGATCTGAGATCTGAGATCTGAG 720
QY 881 GCAGCCGATGCTGAGAGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGG 940
DB 721 GCAGCCGATGCTGAGAGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGG 780
QY 941 TTCCAGAGCTGAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1000
DB 781 TTCCAGAGCTGAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 1001 CCTGAGAGCTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1060
DB 841 CCTGAGAGCTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 1061 GCTGAAAGGCTGCAAGGCAAGGATCCTCTGACCCCTGAGCCGCTGCAAGAGCCGATG 1120
DB 901 GCTGAAAGGCTGCAAGGCAAGGATCCTCTGACCCCTGAGCCGCTGCAAGAGCCGATG 960
QY 1121 CCTTTAGCCCAACCCGATCCATCTCTGAGAGCTGCTGAGAGCTGCTGAGAGCTG 1180
DB 961 CCTTTAGCCCAACCCGATCCATCTCTGAGAGCTGCTGAGAGCTGCTGAGAGCTG 1020
QY 1181 GGGCTGAGCAAGGCAAGGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1237
DB 1021 GGGCTGAGCAAGGCAAGGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1077

```

```

RESULT 7
AAZ61155
ID AAZ61155 standard; cDNA; 1085 BP.
XX
AC AAZ61155;
XX
DT 30-MAY-2000 (first entry)
XX
DE cDNA J0503-KS encoding domains VIA to XI of a protein kinase.
XX

```



```

XX Key Location/Qualifiers
FH CDS 1..1083
FT /tag= a
FT /product= "htrb-3"
FT /note= "ribbles polypeptide"
XX
XX MO200253743-A2.
XX
XX 11-JUL-2002.
XX
XX 08-JAN-2002; 2002WO-US000070.
XX
XX 08-JAN-2001; 2001US-0260294P.
XX
XX (INTE-) INTERLEUKIN GENETICS INC.
XX
XX Dower S, Qunastrom E, Kiss-Toth E;
XX
XX WPI: 2002-590635/63.
XX
XX P-PSDB; ABB80976.
XX
XX Novel isolated human tribbles homologue-1 polypeptide for inhibiting AP-1
XX mediated inflammatory signal in a cell, and activating ERK-mediated
XX signal e.g. AP-1-mediated gene activation signal in a cell.
XX
XX Example 10; Fig 11A; 13pp; English.
XX
XX The invention provides an isolated human tribbles homologue-1 (htrb-1,
XX also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
XX htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory
XX signal in a cell. The polypeptide employed in the method is preferably
XX htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N htrb-3 C, or htrb
XX -3 N C. It is also useful for providing htrb agonist activity for
XX activating an ERK-mediated signal e.g. AP-1-mediated gene activation
XX signal, an estrogen receptor-mediated gene activation signal, an
XX fibroblast growth factor (FGF) induced signal, or a PMA induced signal,
XX in a cell. Htrb modulators are useful for modulating AP-1 mediated
XX inflammatory signal in a cell such as tumor necrosis factor (TNF) induced
XX inflammatory signal, or an interleukin induced inflammatory signal. htrb
XX proteins are useful in screening assays, predictive medicine and in
XX therapeutics or prophylactics. The htrb proteins are useful for screening
XX compounds e.g. for treating and/or preventing diseases caused by abnormal
XX htrb activity, such as rheumatoid arthritis, diabetes, psoriasis,
XX osteoporosis, diabetic retinopathy, myocardial infarction and cancers.
XX The htrb therapeutics are useful for antagonizing interleukin-1 dependent
XX disorders of human placenta, intraventricular hemorrhage, neonatal white
XX matter damage and subsequent cerebral palsy; and inflammation or
XX autoimmune disorders. The present sequence represents a DNA encoding the
XX htrb-3 polypeptide
XX
XX Sequence 1083 BP; 176 A; 374 C; 327 G; 206 T; 0 U; 0 Other;
XX
Query Match 49.8%; Score 1053; DB 6; Length 1083;
Best Local Similarity 99.0%; Pred. No. 1.4e-254;
Matches 1072; Conservative 0; Mismatches 5; Indels 6; Gaps 1;

```

```

QY 401 CGGGCTTACAGGCGCTGCACTGCTCTTACAGGCACTGAGTATATCTGCAAGGTGTACCCC 460
DB 241 CGGGCTTACAGGCGCTGCACTGCTCTTACAGGCACTGAGTATATCTGCAAGGTGTACCCC 300
QY 461 GTTCAGGAAGCCCTGGCGCTGTGAGAGCCCTTACGGCGGCTGCCCCGACAGCATGTG 520
DB 301 GTTCAGGAAGCCCTGGCGCTGTGAGAGCCCTTACGGCGGCTGCCCCGACAGCATGTG 360
QY 521 GCTCGGCCCATGAGGTCTGTGCTGTGATCCCAAGCTCTCTTACAGGCTTGTACCTGGACC 580
DB 361 GCTCGGCCCATGAGGTCTGTGCTGTGATCCCAAGCTCTCTTACAGGCTTGTACCTGGACC 420
QY 581 CATGGGACATGCACAGCTGTGTGCAAGCGCCACCGATCCCTTACAGGCTGTGCTGCC 640
DB 421 CATGGGACATGCACAGCTGTGTGCAAGCGCCACCGATCCCTTACAGGCTGTGCTGCC 480
QY 641 GTGCTTCTCCGCAATGAGCCACCGCCCTGGGCACTGTACAGGACGCTGTGCTG 700
DB 481 GTGCTTCTCCGCAATGAGCCACCGCCCTGGGCACTGTACAGGACGCTGTGCTG 540
QY 701 CGTGATCTCAAGCTGTGCTGCTTTGCTTCG-----TGACGGTGAAGGAAGAGCTG 754
DB 541 CGTGATCTCAAGCTGTGCTGCTTTGCTTCG-----TGACGGTGAAGGAAGAGCTG 600
QY 755 GTGCTGAGACCTTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 814
DB 601 GTGCTGAGACCTTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 815 AAGCAGCGGTGCGCCAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTG 874
DB 661 AAGCAGCGGTGCGCCAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTG 720
QY 875 GGCAGAGCAGCGGATGTGTGAGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTG 934
DB 721 GGCAGAGCAGCGGATGTGTGAGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTG 780
QY 935 TACCCCTTCAGAGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 994
DB 781 TACCCCTTCAGAGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 995 GCTTGTGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1054
DB 841 GCTTGTGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 1055 GAGCAGCTGAAGCGCTGCAAGCAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1114
DB 901 GAGCAGCTGAAGCGCTGCAAGCAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 960
QY 1115 CGATGCTTGAAGCCCAAGCCGATCCATCTTGTGGAAGGCTGCCAGGTGTGCTGTAT 1174
DB 961 CGATGCTTGAAGCCCAAGCCGATCCATCTTGTGGAAGGCTGCCAGGTGTGCTGTAT 1020
QY 1175 GGACTGGGGCTGAGAGAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1234
DB 1021 GGACTGGGGCTGAGAGAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1235 TAG 1237
DB 1081 TAG 1083

```

RESULT 9
 AAH76218
 ID AAH76218 standard; cDNA, 972 BP.
 XX
 XX AAH76218;
 XX
 XX 29-OCT-2001 (First entry)
 XX
 XX Human kinase PKIN-10 encoding cDNA.
 XX
 XX PKIN; kinase; cytosolic; immunosuppressive; immunostimulant; human;
 XX antiarteriosclerotic; cardiant; gene therapy; antisense therapy; ss.

KW	infertility; cardiovascular disease; coagulation disease; hypertension; ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration; diabetes; anxiety; depression; schizophrenia; viral disease; stroke; gastric ulcer; Alzheimer's disease; gene; ss.
OS	Homo sapiens.
XX	
XX	WO200206315-A2.
XX	
PD	24-JAN-2002.
XX	
PF	17-JUL-2001; 2001WO-11000653.
XX	
XX	18-JUL-2000; 2000IL-00137345.
PR	15-DEC-2000; 2000IL-00140354.
XX	
PA	(COMP-) COMPUGEN LTD.
PI	
XX	Mintz L, Freilich S, Bernstein J;
XX	
DR	WPI, 2002-155037/20.
XX	P-PSDB; ABB06093.
XX	
PT	One hundred and twenty eight novel nucleic acid sequences, useful for treating and diagnosing e.g. cancer, asthma and Alzheimer's.
XX	
PS	Claim 1; Page 107; 290pp; English.
XX	
CC	AB139691 to AB139618 represent novel human nucleic acid sequences encoding the proteins given in ABB06037 to ABB06164. The novel sequences (NS) can have cytoprotective, osteoprotic, gynaecological, neuroprotective, antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide, vasotropic, antiarteriosclerotic, antiinflammatory, dermatological, anorectic, muscular, anti-HIV, antiferility, cardiovascular, anticoncugulant, antifibrinolytic, hypotension, antiasthmatic, cardiant, immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer, antidepressant, gastrointestinal, antileptic, cerebroprotective, neurotropic and contraceptive activities. The NS can be used in vaccines, gene therapy and antisense therapy. Nucleic acids, expression vectors and antibodies from the present invention can be used for treating and diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative diseases, dyscoulia, multiple sclerosis, rheumatoid arthritis, psoriasis, cataracts, retinosis, atherosclerosis, inflammation, skin disorders, glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular disease, coagulation disease, ischaemia, hypertension, asthma, immune disease, epilepsy, angina, neurodegeneration, diabetes, anxiety, depression, schizophrenia, viral disease, gastric ulcers, stroke, Alzheimer's disease and as a contraceptive
XX	
SQ	Sequence 1076 BP; 172 A; 352 C; 346 G; 194 T; 0 U; 12 Other;
	Query Match 33.9%; Score 716.6; DB 6; Length 1076;
	Best Local Similarity 98.0%; Pred. No. 6.1e-170;
	Matches 753; Conservative 3; Mismatches 8; Indels 4; Gaps 3
QY	27 CTAGACCCCGGACAGCGCTGAGCGCTGAGGATCCACAGCTCGGACAGCGACGCG 86
Db	196 CTAGACCCCGGACAGCGCTGAGCGCTGAGGATCCACAGCTCGGACAGCGACGCGA-GAG 254
QY	87 CCGGCGCCACCTGCTGTGCTCTGAGAGCTCTAGAGCCCGGCGGCGCGCCCGACGCGG 146
Db	255 CCGGCGCCACCTGCTGTGCTCTGAGAGGCTCTAGAGCCCGGCGGCGCGCCCGACGCGG 314
QY	147 AACGACGGGGCGAGATGCGAGCGACCCCTCTGGCTGCTCTCGGGTTCCTGTCACAGA 206
Db	315 AACGACGGGGCGAGATGCGAGCGACCCCTCTGGCTGCTCTCGGGTTCCTGTCACAGA 374
QY	207 AGAAGCGGTTGAGTTGATGACAACTTAGATACGAGCGTCCCGTCAGAAACGAGCTC 266
Db	375 AGAAGCGGTTGAGTTGATGACAACTTAGATACGAGCGTCCCGTCAGAAACGAGCTC 434
QY	267 GAAGTGGGCGCCAGCCGACATGCCCCCTGCTGTGGTCCCTGAGCCCACTTACTGCTC 326

Db	435	GAAGTGGGGCCCAAGCCAGACTGCCCCCTGCTGTGGTCCCTGAGCCCACTACTGCTC	49
QY	327	CAGATCGTCAACTGCTGTGGCCACTGAGCCCTCCCGTTTGGGCTCTATGTCCTCTGGAGC	386
Db	495	CAGATCGTCAACTGCTGTGGCCACTGAGCCCTCCCGTTTGGGCTCTATGTCCTCTGGAGC	554
QY	387	CCGAGGAGGGCGGGCGGGCCCTTACAGGCCCTTGACCTCACTACAGGCACTAGATATACCT	446
Db	555	CCGAGGAGGGCGGGCGGGCTTACAGGCCCTTGACCTCACTACAGGCACTAGATATACCT	614
QY	447	GCAAGGTGTACCCCGTCCAGGAAGCCCTGGCGGTGTGAGCCCTTACCGCGGCTGCC	506
Db	615	GCAAGGTGTACCCCGTCCAGGAAGCCCTGGCGGTGTGAGCCCTTATGCGGCTGCC	674
QY	507	GCGACAGAGATGTGGTGGGCCCATGAGTCTCTGGCTGTGATCCGAGCTCTTACGCTT	566
Db	675	GCGACAGAGATGTGGTGGGCCCATGAGTCTCTGGCTGTGATCCGAGCTCTTACGCTT	734
QY	567	TTTTCACTCGAGCCCATGGGAGCATGACAGAGCTGGTGGAGACGGCCACCGTACCGTG	628
Db	735	TTTTCACTCGAGCCCATGGGAGCATGACAGAGCTGGTGGAGACGGCCACCGTACCGTG	794
QY	627	AGCTTGAGGCTCCCGTCTCTTCCGCAGATGAGCCACCGCCCTTGCGCACTGTACACAGC	686
Db	795	AGCTTGAGGCTCCCGTCTCTTCCGCAGATGAGCCACCGCCCTTGCGCACTGTACACAGC	854
QY	687	ACGGTCTGGTCTGGTGATCTCAGAGCTGTGTGCTTCTCTCTGGTGGACCGGAGAGA	746
Db	855	ACGGTCTGGTCTGGTGATCTCAGAGCTGTGTGCTTCTCTCTGGTGGACCGGAGAGA	913
QY	747	AGAACTGTGTCTGGAGAACTGTGAGACTCTCCGCTGCTGACTGGAGC	794
Db	914	AGAACTGTGTCTGGAGAACTGTGAGACTCTCCGCTGCTGACTGGAGC	959

RESULT 11
 ABL39762
 ID ABL39762 standard; cDNA; 1076 BP.
 XX ABL39762;
 AC
 DT 10-MAY-2002 (first entry)
 XX
 DE Human NS cDNA sequence SEQ ID NO:72.
 XX
 KW Human; cytosolic; osteopathic; gynaecological; neuroprotective;
 KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
 KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
 KW anorectic; muscular; antiferility; cardiovascular; anticoagulant;
 KW antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiac;
 KW anticonvulsant; antidiabetic; tranquiliser; immunomodulator; neuroleptic;
 KW gastrointestinal; vinuclide; antitumor; cerebroprotective; noctropic;
 KW contractile; vaccine; gene therapy; cancer; osteoporosis; dystonia;
 KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
 KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KW ischaemia; asthma; immune disease; coagulation disease; hypertension;
 KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KW gastric ulcer; Alzheimer's disease; gene; ss.
 XX
 OS Homo sapiens.
 PN WO200206315-A2.
 XX
 PD 24-JAN-2002.
 XX
 PF 17-JUL-2001; 2001WO-IL000653.
 XX
 PR 18-JUL-2000; 2000IL-00137345.
 PR 15-DEC-2000; 2000IL-00140354.
 XX
 PA (COMP-1) COMPUGEN LTD.

CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 2004 BP; 410 A; 563 C; 590 G; 441 T; 0 U; 0 Other;
 SQ

Query Match 29.4%; Score 621.6; DB 9; Length 2004;
 Best Local Similarity 71.9%; Pred. No. 6,4e-146;
 Matches 880; Conservative 0; Mismatches 314; Indels 30; Gaps 4;

QY 129 GGGCGCGGGGCCCAAGCGGAGACGAGCGGCGAGATGCGAGCCACCCCTCTGGCTCTCTG 188
 Db 122 GGGCGAGGGGCGCTCTGAGACCTGAGACAAATGCGAGCCACATCTCTGGCTCTCTG 181
 QY 189 CGGGTTCCCTGTCTCAGAGAAAGAGCGGTTGAGTTGATGACACTTAAGATACCGAGGTC 248
 Db 182 CGGATTTCCCTCAGAGAAAGAAACCGTTGAGTTTATGACAACTTGAATGTCAGATGTC 241
 QY 249 CGGTCCAGAAACGAGCTCGAAGTGGGCCCGAGCCCAAGACTGCCCCCTGCTGTTGCCCC 308
 Db 242 CAGTCTTAAACGAGTGAAGAGATGAGCTGAGCTTGAGCCACACCC----- 288
 QY 309 TGAGGCCACCTAAGTCTCGAGATCGTCAACTGCTGTGGCACTGCTCCCTCTTGGGC 368
 Db 289 --AGCTACCTCCCGCTCAGACTTGTCACTGCTGTGGCCCTCGCAACCCGGCTGGGC 346
 QY 369 CCTATGCTCTCTGAGGCCCGAGAGGGGGGGGCTTACAGAGCCCTGCACTGGCCTA 428
 Db 347 CCTATATCTCTTTGGAGACGAGAGAGCACTGCACTTATGAGGCCCTGCACTGGCCTA 406
 QY 429 CAGGCACTAGTATACCTGTCAGAGTGTACCCGCTCCAGAGAGCCCTGAGCTGTGAGC 488
 Db 407 CAGGCACTAGTATACCTGTCAGAGTGTATCCCGCCAGAGGCCGAGGGGCTGTGGCAGC 466
 QY 489 CCTAGCGCGGCGTCCCCCGGCAAGCATGTGGCTGTGGCCACTGAGTCTGTGGCTGTGA 548
 Db 467 CGTATGCAAGACTGCTTACCAACGACAGGTGGCCCGTCCACAGAGGCTCTGTGGGCT 526
 QY 549 CCAGCTCTCTACGCTTTTTCACCTCGAGCCATGAGGAGCAATGACAGCTGTGGCAG 608
 Db 527 CTGAGTCTCTATACACTTTTTCAGAGAAACCAATGAGGAGCACTTGTGAGCGCA 586
 QY 609 CGGCGCACCGTATCCCTTGAAGCTGAGCTGCGTGTCTTCCGCGAGATGGCCACCGCC 668
 Db 587 GCGCGCGGGTATCCCGGAGCCCGAGGCTGCGCGCTTTTCGAGAGATGGCTATGGCG 646
 QY 669 TGGCGCATGTACACGACAGCAGCTGTGTCTGCTGCTGATTTCAAGCTGTGTGCTTGTCT 728
 Db 647 TGGCACTGCGACAGACAGCGGCTCATCTTTCGAGACTCAAGTTGCTGATTTGCT 706
 QY 729 TCGTGAACCGTGAAGAGAAAGAGTGTGTGAGAACTTGAGGACTGTCGCTGA 788
 Db 707 TCAGCAACTGTGAAGAGAGAGAGTGTGTGAGAACTTGAGGAGTGTGTGATGA 766
 QY 789 CTGGGCGCAGATGATTCCTGTGTGAGCAAGACGCGTGCACCTTACGTGGAGACCTGAGA 848
 Db 767 CTGAGCCAGACGACTCTGTGTGAGCAAGACGCGTGTGCTATGTGGAGCAAGAGA 826
 QY 849 TACTCAGTCAACGGGCTCATATCTGGGGCAAGGCAAGCTGTCTGAGAGCTGTGGCGTGG 908
 Db 827 TACTCAGTCCCGGCTCATATCTCGGCAAGGCGGCTGATGTCTGAGAGCTGTGGCGTGG 886
 QY 909 CGCTTTTCAAGATGTGGCGCGGCTCACTACCCCTTCAGAGACTCGGAGGCTGTCCGCTCT 968
 Db 887 CGCTTTTCAAGATGTGGCGCGGCTCACTACCCCTTCAGAGACTCGGAGGCTGTCCGCTCT 946
 QY 969 TCGGCAAGATCGCGCGGCGCTTACGCTTTCAGAGGCTCTCGGCGCTCTGCGCCCT 1028
 Db 947 TTGGCAAGATCGGAGAGGAGACCTTTGCTGCTGAGGAGGCTTACGCTCTGTGTGAT 1006
 QY 1029 GTCTGAGTGGCTGCTCTGTGTGGAGGCAAGCTGAGAGCGCTACAGCCACAGGCAATCC 1088
 Db 1007 GCTTATCGGCTGTCTCTCGCAAGGAGCTTACAGACGACTTGTGGCCCTGAGGAAATCC 1066
 QY 1089 TCTGCAACCCCTGCTGCGACAGACCCGATGCTTGAAGCCCAACCGGATCCCATCTCT 1148

Db 1067 TGCTATCTCTGTGTGAGAGAGAGATTGACAGCCAGCTCTCTCCACGGTCTGACGAA 1126
 QY 1149 GGGAGGCTGCCAGTGTGTCTGATGAGTCTGGAGCAAGACGAGAGAGAGAGG 1208
 Db 1127 GGGAGATGAGACAGGTTGTCTCCAGATGGCCACAGCTGGAGGAGGAGAGAGGAG- 1185
 QY 1209 GAGACAGAGAGTGTCTGTATGCTTAGACCACTT-ACATACAGCTCAGCTGCGCAAC 1267
 Db 1186 -----GTGGGGCTGTATAGGCTTAAAGCCACTTAAGTGGCCCTTCAAGTCAAGG 1234
 QY 1268 AGTGATTTAGTTGGGGTAGTCTCAAGCTTTCTCTGCTGTAATGAGCCAACT 1327
 Db 1235 TGTGAGTTAGTATGTTGTCTCAGCTTTGCGCTGCT--TTTGGGCCAAGCTAACTG 1292
 QY 1328 TCAGTCTTCCAGAGAGGAGAA 1351
 Db 1293 TAAATCTCTTCTGAGAGAGAGAA 1316

RESULT 13
 ADB53749
 ID ADB53749 standard; DNA; 2004 BP.
 XX

AC ADB53749;
 XX
 DT 04-DEC-2003 (first entry)
 XX

DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:4291.
 XX

KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
 KM toxicity marker; toxicity progression; drug screening;
 KW primary rat hepatocyte toxicity modelling; gene; ds.
 XX

OS Rattus norvegicus.
 XX

PN W02003065993-A2.
 XX

PD 14-AUG-2003.
 XX

PF 04-FEB-2003; 2003MO-US003482.
 XX

PR 04-FEB-2002; 2002US-0351171P.
 PR

PR 13-MAR-2002; 2002US-0363534P.
 PR

PR 08-APR-2002; 2002US-0370248P.
 PR

PR 10-APR-2002; 2002US-0371134P.
 PR

PR 10-APR-2002; 2002US-0371135P.
 PR

PR 11-APR-2002; 2002US-0371413P.
 PR

PR 19-APR-2002; 2002US-0373601P.
 PR

PR 22-APR-2002; 2002US-0374139P.
 PR

PR 08-MAY-2002; 2002US-0378370P.
 PR

PR 09-MAY-2002; 2002US-0378652P.
 PR

PR 09-MAY-2002; 2002US-0378653P.
 PR

PR 09-JUL-2002; 2002US-0394230P.
 PR

PR 09-JUL-2002; 2002US-0394253P.
 PR

PR 04-SEP-2002; 2002US-0407688P.
 PR

PR 28-JAN-2003; 2003US-0442900P.
 XX

PA (GENE-) GENE LOGIC INC.
 XX

PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
 PI Elashoff M;
 XX

DR WPI, 2003-731472/69.
 XX

PT Determining if a compound induces a toxic effect on a tissue or cell, for
 PT identifying hepatotoxic compounds, comprises comparing a gene expression
 PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
 PT mean values.
 XX

XX (SUMU) SUMITOMO SEIYAKU KK.
 PA MPI: 2003-601360/57.
 XX P-PSDB; ADCS9335.
 DR
 XX Controlling cell-death by administering positive or negative regulator of
 PT cell-death inhibitory-factor, for treating cancer, AIDS, autoimmune
 PT diseases, Crohn's diseases, glaucoma, Alzheimer disease.
 XX
 PS Claim 1; SEQ ID NO 2; 35pp; Japanese.
 XX
 CC This invention relates to a novel method for controlling cell-death
 CC comprising administering a positive or negative regulator of cell-death
 CC inhibitory-factor. The invention also comprises a method for screening
 CC modulators of cell-death, by contacting cells expressing cell-death
 CC inhibitory factor with a candidate compound, monitoring level of
 CC expression of cell-death inhibitory factor, evaluating cell-death
 CC modulation ability of the compound based on change in the level of
 CC expression of the factor and selecting compounds having cell-death
 CC modulation ability. The cell death regulator of the invention may have
 CC cardiant, neuroprotective, anti-HIV, antiinflammatory, cerebroprotective,
 CC hepatotropic, cytosolic, immunosuppressive, antineumatic,
 CC ophthalmological, nootropic, antiparkinsonia, anticonvulsant,
 CC hypotensive, antiarteriosclerotic, haemostatic, antialcoholic and
 CC virucide activities. The method of the invention is useful for treating
 CC ischemic damage and congestion, disorder of central nervous system,
 CC hepatic disorder, acute pancreatic inflammation, and cancer, AIDS,
 CC autoimmune diseases, rheumatism, Crohn's diseases, glaucoma, Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, myocardioc infarction, cerebral
 CC trauma, cerebral infarction, reperfusion damage, myocardial infarction, cerebral
 CC hepatitis, and cerebral ischemia. The present sequence represents the DNA
 CC sequence encoding the rat cell death inhibitory protein of the invention.
 XX
 SQ Sequence 2004 BP; 410 A; 563 C; 590 G; 441 T; 0 U; 0 Other;

Query Match 29.4%; Score 621.6; DB 9; Length 2004;
 Best Local Similarity 71.9%; Pred. No. 6,4e-146;
 Matches 880; Conservative 0; Mismatches 314; Indels 30; Gaps 4;

QY 129 GGGCCCGGGCCACGCGGAGAGATGCGAGCCACCCCTGCTGCTCTG 188
 DB 122 GGGCCCGGGCCCTGAGACCTCAGGACAAAGATGAGCCACTCTGCTGCTCTG 181
 QY 189 CGGCTTCCCTGTCAGAGAGAGCGTTGAGTTGATGACAACTTAGTACGAGCGTC 248
 DB 182 CGGATGTTCCCTGCAAGAGAGAACCGTTGAGTTGATGAGCAACATTATGTCGAGTCTC 241
 QY 249 CGGTCCAGAAAGAGACTGGAAGTGGGCCCGGACCAAGCTGCCCCCTGCTGTTGCCCC 308
 DB 242 CAGTCTTAAAGAGATGAGATGAGCTGAGCCCTGAGCAACACCC----- 288
 QY 309 TGAGCCCACTACTCTCCAGATCGTGCACTGCTGAGCCACTCCCTCCCTTTGGGC 368
 DB 289 --AGCTTACTCCCGCTCAAGCTTGTCACTGCTGTGCCCCCTGCAACCCGGCTGGGCG 346
 QY 369 CCTATGTCCTCTGAGGCCGAGAGAGCGCGGCTTACAGGCGCTGCACTGACCTTA 428
 DB 347 CCTATATCTCTTTGGAAGAGAGCAAGGCACTGCACTATGCGGCCCTGCACTGCCCCA 406
 QY 429 CAGGACCTGAGTATATCTGCAAGGTGTACCCCGTTCAGAGAAAGCCCTGCGCTGTGAGC 488
 DB 407 CAGGACAGAGATACCTGCAAGGTGTATCCCGCAGAGAGGCCCGAGGCTGTGCGAC 466
 QY 489 CCTACGCGCGGCTGCCCCCGCAGCAAGATGTGGCTCGGCCCACTGAGGTCTGTGGTGTGA 548
 DB 467 CGTATGCAAGACTGCTTACCAACAGACAGTGGCCCGTCCCAAGAGGTCTGTGGTGGCT 526
 QY 549 CCGAGCTCTCTAGCGCTTTTCACTCGAACCATGAGGACATGACAGCGCTGTGCGAA 608
 DB 527 CTCAGCTCTTATAGACCTTTTTCAGAGAGACCCATGGGGACTTGGACAGTCTGGGTGGCA 586

QY 609 CGGCGCAACGTAATCCCTGAGCCTGAGGCTGCGCTGCTTCCGCAAGATGGCCACGCGCC 668
 DB 587 GCCCGCGGTAATCCCGAGCCGAGCGGCTGCGCGCTTTTCCAGAGATGGCTAGTGGCG 646
 QY 669 TGGCGCACTGTACCAACGACAGCGTGTGCTCTGCTGATGTCAAGCTGTGTGCTTGTCT 728
 DB 647 TGGCAACATGCCACAGAGACAGCGGCTCATCTTGGCGACCTCAAGTTGGCGATTGTCT 706
 QY 729 TCGCTGACCGTGAAGAGAAAGCTGTGTAAGAACTGGAGACTCTGCTGTGA 788
 DB 707 TAGCAACTGTGAGAGAGAAAGTGTGTCTGAGAACTGGAGATGCTGTGTATGA 766
 QY 789 CTGAGCCAGATGATTCCTGTGAGCAAGACAGCGGCCAGCTTCGTGGAGCTTGGAG 848
 DB 767 CTGAGCAACAGACTCTCTGTGGAGCAACAGCGGCTGCTTATGTGGAGCAAGAGA 826
 QY 849 TACTCAGCTCAAGGCGCTTATATCTGGGCAAGGACCGATGTGTGAGCTGGGCGTGG 908
 DB 827 TACTCAGCTCCCGGCATCTCTCCGAGAGCGGCTGATGTGTGAGCTGGGCGTGG 886
 QY 909 CGCTCTTACCAAGCTGGCGCGGCACTACCCCTTCAGAGACTTGGAGCTGTCTCTCT 968
 DB 887 CGCTCTTACCAAGCTGGCGGCTGAGTACCGTTCAGAGACTTGTGAACAGCCCTGCTCT 946
 QY 969 TGGGCAAGATCCGCGCGGCGCTTACGCTTGTGCAAGGCTCTGCGCCCGTGGCGCT 1028
 DB 947 TTGGCAAGATCCGTAAGAGGACCTTTGGCCCTGCTGAGGCGCTTATAGCTTGTCTGAT 1006
 QY 1029 GTCTGCTGTGCTGCTCTCTCTGTCGAGAGCCAGTGAACGAGTCAAGCCACAGCATCC 1088
 DB 1007 GCCTATCCGCTGTCTCTCTCCGAGAGGCTTCAAGAGCACTTGTGCGCTGGAGATCC 1066
 QY 1089 TCGTCAACCCCTGGCTGGAGACAGACCGGATGCTTGAACCCAGTCCCATCTCT 1148
 DB 1067 TCTTCAATCCCTGTGTGAGAGAGATGCAAGCAAGTCTTCTTCCACGCTTGAACGAA 1126
 QY 1149 GGGAGGCTGCGCAGGCTGCTCCCTGATGACTGAGGCGCTGAGCAAGAGCAAGAGAGAG 1208
 DB 1127 GGGAGATGAGCAGGCTTGTCCAGATGAGGCGCAGAGCTGAGAGGCAAGAGAGAGAG- 1185
 QY 1209 GAGACAGAAAGTGTCTGTATGAGTGAACCACTT-ACTTACAGCTCACTGCTCAAC 1267
 DB 1186 -----GTGGGCTGTATGAGCTTGAAGCCCACTTAACTGAGCCCTTCACTCAAGG 1234
 QY 1268 AGTGATGAGTTGGGGGTAGCTCCAGGCTTCTCCCTGCTTGAACCTGAGCCAACT 1327
 DB 1235 TGTGATGAGTACTGAGTGTCTCAAGCTTGTGCGCTGCT--TTTGGGCCAAGCTTAACTG 1292
 QY 1328 TCAGTCCCTTCCAGAGGAGAGAA 1351
 DB 1293 TAAAGCTCTTCTGAGAGAGAGAA 1316

RESULT 15
 AAS91231/c
 ID AAS91231 standard; cDNA; 996 BP.
 XX AAS91231;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #27035.
 XX
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 XX
 XX W0200175067-A2.
 XX
 XX PD 11-OCT-2001.
 XX

OM nucleic - nucleic search, using sw model

Run on: August 29, 2004, 04:49:36 ; Search time 8342 Seconds

(without alignments)
10994.229 Million cell updates/sec

Title: US-10-070-337-16
Perfact edoxo: 3116

Sequence: 1 ggagcgcgctccgcgcgct.....tatgaataaaaaaaaaa 2116

Scoring table: IDENTITY_NUC

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match	0%
--------------------------------	----

Listing first 45 summaries

Database :

```

1:  gb ba:*
2:  gb htg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb pl:*
9:  gb pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em ba:*
16: em fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em htg_hum:*
31: em htg_inv:*
32: em htg_other:*
33: em htg_mus:*
34: em htg_pln:*
35: em htg_rod:*
36: em htg_nam:*
37: em htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	216	100.0	2116	6	AX0993934	AX0993934 Sequence
2	2100.2	99.3	2389	6	AR411699	AR411699 Sequence
3	2095.4	99.0	2257	9	AK026945	AK026945 Homo sapi
4	2093.8	99.0	2283	9	BC022484	BC022484 Homo sapi
5	2078.2	98.2	2488	9	AY247738	AY247738 Homo sapi
6	2046.2	96.7	2121	9	BC019363	BC019363 Homo sapi
7	1367	64.6	153170	9	HS110367	HS110367 Human DNA
8	1072.2	50.7	1077	6	AX166518	AX166518 Sequence
9	1070.6	50.6	1077	12	AY335754	AY335754 Synthetic
10	1069.8	50.6	1085	6	BD248318	BD248318 Human cDN
11	1069.2	50.5	1074	6	AR411700	AR411700 Sequence
12	1053	49.8	1083	6	AX572896	AX572896 Sequence
13	1053	49.8	1083	9	AF250311	AF250311 Homo sapi
14	803.8	38.0	972	6	AX224734	AX224734 Sequence
15	716.6	33.9	1076	6	AX364906	AX364906 Sequence
16	716.6	33.9	1076	6	AX364921	AX364921 Sequence
17	670.6	31.7	1280	10	MM0514260	MM0514260 Mus muscu
18	656.6	31.0	1256	10	AF388868	AF388868 Mus muscu
19	621.6	29.4	1204	10	AB020967	AB020967 Rattus sp
20	533.8	25.2	1969	6	BC012955	BC012955 Mus muscu
21	533	25.2	541	6	AX381642	AX381642 Sequence
22	505.2	23.9	2259	6	AR274998	AR274998 Sequence
23	505.2	23.9	2559	6	HS002019	HS002019 Human AU-Ri
24	505.2	23.9	2562	6	T24432	T24432 Sequence 4
25	499.2	23.6	75131	2	AC023022	AC023022 Homo sapi
26	354.2	16.7	336	6	AX245889	AX245889 Sequence
27	318	15.0	98566	10	AL928568	AL928568 Mouse DNA
28	315.4	14.9	353	6	AX329832	AX329832 Sequence
29	295.6	14.0	235947	2	AC133773	AC133773 Rattus no
30	295.6	14.0	282849	2	AC132555	AC132555 Rattus no
31	293.8	13.9	75131	2	AC023022	AC023022 Homo sapi
32	278.2	13.1	331	11	G23310	G23310 human STS w
33	269.4	12.7	3319	6	AX572894	AX572894 Sequence
34	269.4	12.7	3319	9	AF250310	AF250310 Homo sapi
35	267.8	12.7	3317	9	AF205437	AF205437 Homo sapi
36	267.8	12.7	3338	9	HS003656	HS003656 Homo sapi
37	267.8	12.7	3645	9	BC063292	BC063292 Homo sapi
38	256.6	12.1	2514	10	BC006800	BC006800 Mus muscu
39	253.4	12.0	1244	10	AF358866	AF358866 Mus muscu
40	252.4	11.9	1682	4	AY247741	AY247741 Bos tauru
41	251.4	11.9	1032	12	AY335660	AY335660 Synthetic
42	251.4	11.9	2147	9	BC002637	BC002637 Homo sapi
43	251.4	11.9	2156	9	AX779968	AX779968 Sequence
44	251.4	11.9	3062	9	AY245544	AY245544 Homo sapi
45	251.4	11.9	4221	9	D87119	D87119 Homo sapien

ALIGNMENTS

RESULT 1	AX099934	2116 bp	DNA	linear	PAT 02-APR-2001
LOCUS	AX099934				
DEFINITION	Sequence 16 from Patent WO0120004.				
ACCESSION	AX099934				
VERSION	AX099934.1	GI:13518944			
KEYWORDS	.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	1				
AUTHORS	Yue, H., Tang, Y. T., Bandman, O., Hillman, J. L., Baughn, M. R.,				
TITLE	Arimai, Y. and Lu, D. A. Protein phosphatase and kinase proteins				

JOURNAL Patent: WO 0120004-A 16 22-MAR-2001;
Incyte Genomics, Inc. (US)
Location/Qualifiers
FEATURES
source 1. 2116
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID NO: 1271505CB1"

ORIGIN

Query Match 100.0%; Score 2116; DB 6; Length 2116;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GGAGGCGGCTCCGCGCGGCTCCGCTGTAGAACCCGGGCGAGGCTGAGCTGGAGCTGGAG 60
Db 1 GGAGGCGGCTCCGCGCGGCTCCGCTGTAGAACCCGGGCGAGGCTGAGCTGGAGCTGGAG 60
QY 61 TCCCGAGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 61 TCCCGAGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 121 CCCCCGCGCGCCCGGCGCCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 121 CCCCCGCGCGCCCGGCGCCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 181 TGTCTCTGCGGGTTCCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 181 TGTCTCTGCGGGTTCCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 CGAGCGTCCCGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 241 CGAGCGTCCCGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 GTTGCCCTGAGGCGCACTACTGCTCCAGATCTGTGCACTGTGTGCACTGTGTGCACT 360
Db 301 GTTGCCCTGAGGCGCACTACTGCTCCAGATCTGTGCACTGTGTGCACTGTGTGCACT 360
QY 361 TCTTGGGCGCTATGTCTCTCTGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 TCTTGGGCGCTATGTCTCTCTGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 CTGCGCTTACAGGAGCTGAGATATCTTGAAAGGTGACCCCTCAAGAGAGAGAGAGAG 480
Db 421 CTGCGCTTACAGGAGCTGAGATATCTTGAAAGGTGACCCCTCAAGAGAGAGAGAGAG 480
QY 481 GCTGAGAGCCCTTACGCGCGGCTGCGCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 GCTGAGAGCCCTTACGCGCGGCTGCGCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 GGTCTGTACCAAGTCTCTTACGCGCTTTTCACTCGGAGCCCATGGGAGATGCAAGCT 600
Db 541 GGTCTGTACCAAGTCTCTTACGCGCTTTTCACTCGGAGCCCATGGGAGATGCAAGCT 600
QY 601 GGTGCGAAGCGGCAACGCTATCCCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 660
Db 601 GGTGCGAAGCGGCAACGCTATCCCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 660
QY 661 CACCGCGCTGCGGCACTGTCAACGAGCAGGCTGTGTCTGTGATCTCAAGCTGTGTCT 720
Db 661 CACCGCGCTGCGGCACTGTCAACGAGCAGGCTGTGTCTGTGATCTCAAGCTGTGTCT 720
QY 721 CTTTGTCTTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 CTTTGTCTTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 CGTGTGACTGGGCGAGATGTTCCCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 CGTGTGACTGGGCGAGATGTTCCCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 ACCTGAGATACTAGCTCAGCGGCGCTCATCTCGGGGAGAGAGAGAGAGAGAGAGAGCT 900
Db 841 ACCTGAGATACTAGCTCAGCGGCGCTCATCTCGGGGAGAGAGAGAGAGAGAGAGAGCT 900
```

```
QY 901 GGGCGTGGGCGCTCTTACCAATGCTGGCGGCGCACTACCCCTTCCAGAGACTGGAGCTGT 960
Db 901 GGGCGTGGGCGCTCTTACCAATGCTGGCGGCGCACTACCCCTTCCAGAGACTGGAGCTGT 960
QY 961 CTTGCTCTTGGCAAGATCCGCGCGGCGCTTACGCTTGTCTGAGAGGCTTCTGGCCCC 1020
Db 961 CTTGCTCTTGGCAAGATCCGCGCGGCGCTTACGCTTGTCTGAGAGGCTTCTGGCCCC 1020
QY 1021 TCCCGCTGTCTGTGTGCTGCTCTCTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TCCCGCTGTCTGTGTGCTGCTCTCTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 AGGCAATCCCTCGACACCCCTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1140
Db 1081 AGGCAATCCCTCGACACCCCTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1140
QY 1141 CCATCTCTGGAGAGCTGCGCAGAGTGTCCCTGATGAGCTGGAGGCTGAGAGAGAGAG 1200
Db 1141 CCATCTCTGGAGAGCTGCGCAGAGTGTCCCTGATGAGCTGGAGGCTGAGAGAGAGAG 1200
QY 1201 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 1201 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 TGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 TGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 CAAACCTTCAAGTCCCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 CAAACCTTCAAGTCCCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 AACATCTGCTTGTGTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 AACATCTGCTTGTGTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 CCCTGTTCTCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1441 CCCTGTTCTCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 GAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 1501 GAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1561 TACACTGGGTACACTTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1561 TACACTGGGTACACTTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 TGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1621 TGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1681 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1681 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
QY 1741 AAGCCAGCTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1800
Db 1741 AAGCCAGCTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1800
QY 1801 CTCCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1860
Db 1801 CTCCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1860
QY 1861 GTTCTGTCTTAACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 1861 GTTCTGTCTTAACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
QY 1921 TCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1980
Db 1921 TCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1980
```


QY 1981 GGAAGTCCAGGTGGAGCTCTTCTGAGGACACTTGGGGTCCAAATCCAGGTCATAC 2040
Db 1981 GGAAGTCCAGGTGGAGCTCTTCTGAGGACACTTGGGGTCCAAATCCAGGTCATAC 2040
QY 2041 TCTAGTTTGTGATACCATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
Db 2041 TCTAGTTTGTGATACCATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
QY 2101 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2116
Db 2101 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2116

RESULT 2
AR411699 2389 bp DNA linear PAT 18-DEC-2003
LOCUS AR411699
DEFINITION Sequence 7 from patent US 6638721.
ACCESSION AR411699
VERSION AR411699.1 GI:40164080
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2389)
AUTHORS Meyers, R., Kapeller-Libermann, R. and Williamson, M.
TITLE Human protein kinases and uses therefor
JOURNAL Patent: US 6638721-A 7 28-OCT-2003;
FEATURES
Location/Qualifiers
1..2389
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 99.3%; Score 2100.2; DB 6; Length 2389;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2113; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 GGAAGGCTCCGCGGCTCCGCTGCTGCTAGAGCCCGGGCAGGCTGAGCTGGGCTGGGA 60
Db 222 GGAAGGCTCCGCGGCTCCGCTGCTGCTAGAGCCCGGGCAGGCTGAGCTGGGCTGGGA 281
QY 61 TCCGAGCTCCGCGAGCAGCAGCA-GCGGCGGCGCACTGCTGCTGCTGCTGCTGCTGCTG 119
Db 282 TCCGAGCTCCGCGAGCAGCAGCA-GCGGCGGCGCACTGCTGCTGCTGCTGCTGCTGCTG 341
QY 120 GCG 179
Db 342 GCG 401
QY 180 CTGCTCCGCGGCTTCT 239
Db 402 CTGCTCCGCGGCTTCT 461
QY 240 CCGAGCGTCCCGTCCAGAAAGAGCTGAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 299
Db 462 CCGAGCGTCCCGTCCAGAAAGAGCTGAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 521
QY 300 TGTGGCGCTGAGCGCACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 359
Db 522 TGTGGCGCTGAGCGCACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 581
QY 360 GTCTTGGGCGCTTGTCT 419
Db 582 GTCTTGGGCGCTTGTCT 641
QY 420 ACTGCGCTTACAGGCACTGATATCTGCAAGGTGTACCGCGTCCAGAAAGCGCTGGCG 479
Db 642 ACTGCGCTTACAGGCACTGATATCTGCAAGGTGTACCGCGTCCAGAAAGCGCTGGCG 701
QY 480 TGTGAGAGCGCTTACAGGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 539
Db 702 TGTGAGAGCGCTTACAGGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 761

QY 540 TGGCTGTATCCAGGCTCCTCTACAGGCTTTTTCAGCTGAGCCCATGGGAGCATCAGAGCC 599
Db 762 TGGCTGTATCCAGGCTCCTCTACAGGCTTTTTCAGCTGAGCCCATGGGAGCATCAGAGCC 821
QY 600 TGGTGGAAAGCGGCGCACTGATCTCTGAGCTGAGGCTGCGCTCTTCCGCGAGATG 659
Db 822 TGGTGGAAAGCGGCGCACTGATCTCTGAGCTGAGGCTGCGCTCTTCCGCGAGATG 881
QY 660 CCAGCGCGCTGGCGGAGCTGTCACAGCAAGCTGTCTCTGCTGATCTCAAGCTGTGTC 719
Db 882 CCAGCGCGCTGGCGGAGCTGTCACAGCAAGCTGTCTCTGCTGATCTCAAGCTGTGTC 941
QY 720 GCTTGTCTTCTGCTGAGCAGTGAAGAAAGCTGCTGCTGAGAACCTGGAGAGCTCT 779
Db 942 GCTTGTCTTCTGCTGAGCAGTGAAGAAAGCTGCTGCTGAGAACCTGGAGAGCTCT 1001
QY 780 GCGTCTGACTGAGCGCAATGATCTCTGAGGAGCAAGCAGCGCTGCGGCTGAGCTG 839
Db 1002 GCGTCTGACTGAGCGCAATGATCTCTGAGGAGCAAGCAGCGCTGCGGCTGAGCTG 1061
QY 840 GACCTGAGATCTCAGCTCAGCGGCTGATCTCGGCGCAAGGCAAGCCGATCTGTGAGCC 899
Db 1062 GACCTGAGATCTCAGCTCAGCGGCTGATCTCGGCGCAAGGCAAGCCGATCTGTGAGCC 1121
QY 900 TGGGCGTGGCGCTCTTCAACAATGCTGCGCGGCACTACCCCTTCCAGAGCTGAGAGCC 959
Db 1122 TGGGCGTGGCGCTCTTCAACAATGCTGCGCGGCACTACCCCTTCCAGAGCTGAGAGCC 1181
QY 960 TCTGCTCTTGTGGCAAGATCCGCGCGGCGGCTTACGCTTTCAGAGCTTGTGAGCC 1019
Db 1182 TCTGCTCTTGTGGCAAGATCCGCGCGGCGGCTTACGCTTTCAGAGCTTGTGAGCC 1241
QY 1020 CTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
Db 1242 CTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1301
QY 1080 CAGGCACTCTCTGACCCCTGCTGCGAGCAGCCAGATCCCTTACCCCAACCCGAT 1139
Db 1302 CAGGCACTCTCTGACCCCTGCTGCGAGCAGCCAGATCCCTTACCCCAACCCGAT 1361
QY 1140 CCGATCTCTGAGAGGCTGCGCGAGTGTCTCTGATGCACTGGGCTGAGCAAGCAGAG 1199
Db 1362 CCGATCTCTGAGAGGCTGCGCGAGTGTCTCTGATGCACTGGGCTGAGCAAGCAGAG 1421
QY 1200 AAGAGGAGGAGAGAGAGAGTGTCTGATGCTGAGACCACTTACTCAAGCTGAG 1259
Db 1422 AAGAGGAGGAGAGAGAGTGTCTGATGCTGAGACCACTTACTCAAGCTGAG 1481
QY 1260 CTGCGCAAGTGAATGATTTGGGGGTAGCTCCAGAGCTTCTCTGCTCTGAACTGAG 1319
Db 1482 CTGCGCAAGTGAATGATTTGGGGGTAGCTCCAGAGCTTCTCTGCTCTGAACTGAG 1541
QY 1320 CCAAGCTTCACTGCTTCCAGAAAGGAGAAAGCAGAGCTGTGTGAGTGTGCTGTG 1379
Db 1542 CCAAGCTTCACTGCTTCCAGAAAGGAGAAAGCAGAGCTGTGTGAGTGTGCTGTG 1601
QY 1380 TACACATCTGCTTGTGTCACACACATGAGTCTCTGCTGCTTATAGTGTGCA 1439
Db 1602 TACACATCTGCTTGTGTCACACACATGAGTCTCTGCTGCTTATAGTGTGCA 1661
QY 1440 GCCCTGTTCTCGGTGCTGAGAGTACAGCAGTGAAGCAATATTCCTGCTCAC 1499
Db 1662 GCCCTGTTCTCGGTGCTGAGAGTACAGCAGTGAAGCAATATTCCTGCTCAC 1721
QY 1500 AGAGATACAAACTGAGATCTTGAAGTGAACAACCTTTCATGACCATAGTGTACTGT 1559
Db 1722 AGAGATACAAACTGAGATCTTGAAGTGAACAACCTTTCATGACCATAGTGTACTGT 1781
QY 1560 CTACACTGAGGATCACTTTGATACAGTGTGCGGCTTCACTGATCTGTGCTGAGGCACT 1619
Db 1782 CTACACTGAGGATCACTTTGATACAGTGTGCGGCTTCACTGATCTGTGCTGAGGCACT 1841
QY 1620 CTGTCCAGAGCAATCCCTTTCACAAACCAAGCTGCTTGTATCTTGTACTCTTTC 1679

```

Db      1842  CTGTCACAGGCAATCCCTTTCACAAACAAACAGCTGCTTGTATCTTGTACTTTTC 1901
QY      1880  AGAGAAAGGAGATATCCCTGTGCAAGAGCTCCAGGCTCTCCCTGCAACTCAGAGC 1739
Db      1902  AGAGAAAGGAGATATCCCTGTGCAAGAGCTCCAGGCTCTCCCTGCAACTCAGAGC 1961
QY      1740  CAAGCCAGCTCATCTGTGGAACTGTGTCCAGATCTCTGTCTCTTGTATTAAGAT 1799
Db      1962  CAAGCCAGCTCATCTGTGGAACTGTGTCCAGATCTCTGTCTCTTGTATTAAGAT 2021
QY      1800  TCTGCTTCCAGGCTTAAGCTGTGGATTTGGGCGAGATTAAGATTAAGATTAAGCT 1859
Db      2022  TCTGCTTCCAGGCTTAAGCTGTGGATTTGGGCGAGATTAAGATTAAGATTAAGCT 2081
QY      1860  AGTCTTGTCTTAAGCTCAAGACTGTCTGTGAATGAGGCTCCAGGCTGTCAACATG 1919
Db      2082  AGTCTTGTCTTAAGCTCAAGACTGTCTGTGAATGAGGCTCCAGGCTGTCAACATG 2141
QY      1920  TTCTGACCTGACACCAAGGTTGAGGACAGATTAAGGAGGCTGTCTGTGCGCAC 1979
Db      2142  TTCTGACCTGACACCAAGGTTGAGGACAGATTAAGGAGGCTGTCTGTGCGCAC 2201
QY      1980  TGGAAAGTCCCGAGGCTCTCTGTGGGACACTTGGGCTCCAGATCCAGTCCATA 2039
Db      2202  TGGAAAGTCCCGAGGCTCTCTGTGGGACACTTGGGCTCCAGATCCAGTCCATA 2261
QY      2040  CTCTAGTCTTGTGATCAATGATGATGTTTATCTGTCTGCTTAATTAAGAGATTA 2099
Db      2262  CTCTAGTCTTGTGATCAATGATGATGTTTATCTGTCTGCTTAATTAAGAGATTA 2321
QY      2100  GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2116
Db      2322  GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2338

RESULT 3
LOCUS   AK026945      2257 bp      mRNA      linear      PRI 12-SEP-2003
DEFINITION Homo sapiens cDNA: FLJ23292 fls, clone HEP10334.
ACCESSION AK026945
VERSION   AK026945.1 GI:10439923
KEYWORDS  oligo capping; fls (full insert sequence).
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
  Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y.,
  Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,
  Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
  NEDO human cDNA sequencing project
  Unpublished
  2 (bases 1 to 2257)
  Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,
  Shibahara, T., Tanaka, T. and Nakamura, Y.
  Direct Submission
  Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
  University of Tokyo, Laboratory of Genome Structure Analysis, Human
  Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
  Japan (E-mail: f1cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
  Fax:81-3-5449-5416)
COMMENT  NEDO human cDNA sequencing project supported by Ministry of
  International Trade and Industry of Japan. cDNA full insert
  sequencing: Research Association for Biotechnology; cDNA library
  construction, 5'- & 3'-end one pass sequencing: Department of
  Virology and Human Genome Center, Institute of Medical Science,
  University of Tokyo (partly supported by Science and Technology
  Agency).
FEATURES
  source
    1..2257
    /organism="Homo sapiens"
    /mol_type="mRNA"

```

```

CDS
  1..2257
    /db_xref="taxon:9606"
    /clone="HEP10334"
    /cell_line="HePG2"
    /cell_type="hepatoma"
    /clone_id="HEP"
    /note="Cloning vector pWE18FLJ3"
    295..1371
      /note="unlabeled protein product"
      /codon_start=1
      /protein_id="BAB15597.1"
      /db_xref="GI:10439924"
      /translation="MRAPIPLAAPAGSLSRKRLDLNDLTERPVQKARSGPQRLP
      PCLPLSPPTAPDRATVATSRGLGPVLLPEEGGAYVQALHCPDTGTEYCKYVQ
      EAPVLEPPYARLPKHVAPTEVLVAGTOLIVAFTRHGMHSLVSRHRIPEEA
      VLFRQMATLAHGHQVGLVLDLKCREFVADRRKRLVNLHNSLGVLTGDDSLMD
      KHACPAYAPRILSSPASVSGKADVSLGVALFVLMLAGHYPDQSEPVLLFGKIRG
      AYALPAGISAPARCLVRCILRREPARERLTALGILLHPRLQDPMPLATRSHLMEAAQ
      VVPGGLIDEARBEEDREVLYG"

ORIGIN
  Query Match      99.0%; Score 2095.4; DB 9; Length 2257;
  Best Local Similarity 99.7%; Pred. No. 0;
  Matches 2110; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

  1  GAGAGCGGCTCCGCGCGGCTCGCTGTAGAGACCCGGGAGGCTGAGCTGGCTGGGA 60
  Db  GAGAGCGGCTCCGCGCGGCTCGCTGTAGAGACCCGGGAGGCTGAGCTGGCTGGGA 193
  QY  61  TCCGAGCTCGGACAGACCGGA -GGGCGCGGCGACCTGTGTGCTGCTGAGCTGGA 119
  Db  134  TCCGAGCTCGGACAGACCGGA -GGGCGCGGCGACCTGTGTGCTGCTGAGCTGGA 253
  QY  194  TCCGAGCTCGGACAGACCGGA -GGGCGCGGCGACCTGTGTGCTGCTGAGCTGGA 253
  Db  120  GCCCGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 179
  QY  254  GCCCGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 313
  Db  180  CTGCTCTGCGGCTTCCCTGTCTGTCAGAGAAAGACGGTGTGAGTGTGATGACAACTTA 239
  QY  314  CTGCTCTGCGGCTTCCCTGTCTGTCAGAGAAAGACGGTGTGAGTGTGATGACAACTTA 373
  Db  240  CCGAGCGTCCCTCCAGAAAGAGGTCGAAAGTGGGCGGCGGCGGCGGCGGCGGCGG 299
  QY  374  CCGAGCGTCCCTCCAGAAAGAGGTCGAAAGTGGGCGGCGGCGGCGGCGGCGGCGG 433
  Db  300  TGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 359
  QY  434  TGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 493
  Db  360  GTCTTGGGCGGCTATGTCTCTCTGAGAGCCCGGAGAGGCGGCGGCGGCGGCGGCG 419
  QY  494  GTCTTGGGCGGCTATGTCTCTCTGAGAGCCCGGAGAGGCGGCGGCGGCGGCGGCG 553
  Db  420  ACTGCGCTTACAGGACTGAGTATCTGTAAGGTGATCCCTGTCAGAGAGCCCTGGCG 479
  QY  554  ACTGCGCTTACAGGACTGAGTATCTGTAAGGTGATCCCTGTCAGAGAGCCCTGGCG 613
  Db  480  TGTGAGGCGCTTACAGGCGGCTGCCCCGCAAGAGATGTGCTGCGGCCCATAGAGTCC 539
  QY  614  TGTGAGGCGCTTACAGGCGGCTGCCCCGCAAGAGATGTGCTGCGGCCCATAGAGTCC 673
  Db  540  TGTGAGGCGCTTACAGGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 599
  QY  674  TGTGAGGCGCTTACAGGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 733
  Db  600  TGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 659
  QY  734  TGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 793
  Db  660  CCACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 719
  QY  794  CCACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 853
  Db  720  GCTTTGTCTTGTGACCGGTGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 779

```

```

Db      854 GCTTTGCTCTTCGCTGACCGCTGAGAGAGAAAGAGTGTGCTGAGAACTTGGAGAGACTCTCT 913
Qy      780 GGGTGTGACTGGGGCAGATGATTTCCCTGTGGACAAAGACGCGTGCCAGGCTTAAGTGG 839
Db      914 GGGTGTGACTGGGGCAGATGATTTCCCTGTGGACAAAGACGCGTGCCAGGCTTAAGTGG 973
Qy      840 GACCTGAGATCTACGCTACGAGGCTCTATCTCGGGCAAGGACCGCATGTCTGAGACC 899
Db      974 GACCTGAGATCTACGCTACGAGGCTCTATCTCGGGCAAGGACCGCATGTCTGAGACC 1033
Qy      900 TGGGGGTGAGCGCTCTTCAACATGCTGCGGCGCACTACCCCTTCCAGGACTCGGAGCGCTG 959
Db      1034 TGGGGGTGAGCGCTCTTCAACATGCTGCGGCGCACTACCCCTTCCAGGACTCGGAGCGCTG 1093
Qy      960 TCTGTCTTGTGGCAAGATCCGCGGAGGCTTACGCTTGTGCTGAGGCTCTCGGCGCC 1019
Db      1094 TCTGTCTTGTGGCAAGATCCGCGGAGGCTTACGCTTGTGCTGAGGCTCTCGGCGCC 1153
Qy      1020 CTGCGCGGCTGTGCTGTGCTGCTGCTCTTCTGTGCGGAGCCAGCTGAACGGCTCAAGCCA 1079
Db      1154 CTGCGCGGCTGTGCTGTGCTGCTGCTCTTCTGTGCGGAGCCAGCTGAACGGCTCAAGCCA 1213
Qy      1080 CAGGATCTCTCTGACACCCCTGCTGCGGACAGAGACCCGATGCTTACGCTTACGCTTAC 1139
Db      1214 CAGGATCTCTCTGACACCCCTGCTGCGGAGAGAGACCCGATGCTTACGCTTACGCTTAC 1273
Qy      1140 CCGATCTCTGAGAGGCTGCGGAGGCTGCTGCTGCTGATGATGATGATGATGATGATGATGAT 1199
Db      1274 CCGATCTCTGAGAGGCTGCGGAGGCTGCTGCTGCTGATGATGATGATGATGATGATGATGAT 1333
Qy      1200 AAGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1259
Db      1334 AAGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1393
Qy      1260 CTGCAACAGTGTGATTTGAGTTGGGGGAGTAGCTCCAAAGCCTTCTCGGCTCTGAGACTAG 1319
Db      1394 CTGCAACAGTGTGATTTGAGTTGGGGGAGTAGCTCCAAAGCCTTCTCGGCTCTGAGACTAG 1453
Qy      1320 CCAAACTTTCAGTGTGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1379
Db      1454 CCAAACTTTCAGTGTGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1513
Qy      1380 TACACATCTGCTTGTGCTTCCACACACATGCTGCTGCTGCTGCTTATCAGGTGCCAA 1439
Db      1514 TACACATCTGCTTGTGCTTCCACACACATGCTGCTGCTGCTGCTTATCAGGTGCCAA 1573
Qy      1440 GCCCTGTTCTCGGCTGCTGGAGTAGACAGAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1499
Db      1574 GCCCTGTTCTCGGCTGCTGGAGTAGACAGAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1633
Qy      1500 AGAGATGACAAACTGGGATCTCTTGAAGTGAACAACATTTTCCATGACATAGTCACTGT 1559
Db      1634 AGAGATGACAAACTGGGATCTCTTGAAGTGAACAACATTTTCCATGACATAGTCACTGT 1693
Qy      1560 CTACACAGGAGTACACTTGTGACAGAGTGTGCGCTTCACTGATGCTGTGCTCAAGGACCT 1619
Db      1694 CTACACAGGAGTACACTTGTGACAGAGTGTGCGCTTCACTGATGCTGTGCTCAAGGACCT 1753
Qy      1620 CTGTCTCAAGAGCAATCCCTTTCACAAACAAACAGCTGCTTGTGATCTTGACCTTTC 1679
Db      1754 CTGTCTCAAGAGCAATCCCTTTCACAAACAAACAGCTGCTTGTGATCTTGACCTTTC 1813
Qy      1680 AGAGAAAGGAGGATATCCCTGTGCAAAAGGCTCCAGGCTCTCCCTGCAACTCAGAGACC 1739
Db      1814 AGAGAAAGGAGGATATCCCTGTGCAAAAGGCTCCAGGCTCTCCCTGCAACTCAGAGACC 1873
Qy      1740 CAAAGCCAGCTCACTCTGGGAGAGCTGTTCACAGATCTCTGTCTCTTGAATTAAGAGAT 1799
Db      1874 CAAAGCCAGCTCACTCTGGGAGAGCTGTGTTCACAGATCTCTGTCTCTTGAATTAAGAGAT 1933
Qy      1800 TCTCTCTTCAAGCCCTTAAGCCTTGAGATTTGGGCAAGAGATTAAGAACTTAAGAGGCT 1859

```

```

Db      1934 TCTCTTCCAGGCTTAAGCCTTGAGATTGGGGCCAGAGATTAAGAACTTAAGAGGCT 1993
Qy      1860 AGTTCTTGTCTAATCTCAAGACTGTTCTGGAATGAGGGGTCCAGGCTGTCAACCATGAGGC 1919
Db      1994 AGTTCTTGTCTAATCTCAAGACTGTTCTGGAATGAGGGGTCCAGGCTGTCAACCATGAGGC 2053
Qy      1920 TTCTGACCTGAGCACCAGAGGTTGAGGAGCAGAGATTAGGAGGGGTCTGTCTGTGGCACC 1979
Db      2054 TTCTGACCTGAGCACCAGAGGTTGAGGAGCAGAGATTAGGAGGGGTCTGTCTGTGGCACC 2113
Qy      1980 TGGAAAGTCCCAAGGTGGGAGCTTTCTGGGGAGACATTTGGGGTCCACATCCGAGTCCATA 2039
Db      2114 TGGAAAGTCCCAAGGTGGGAGCTTTCTGGGGAGACATTTGGGGTCCACATCCGAGTCCATA 2173
Qy      2040 CTCTAGGTTTGGATACCATAGATGATGATGTTTACCTGTGCTTAATAAGAGAAATTAT 2099
Db      2174 CTCTAGGTTTGGATACCATAGATGATGATGTTTACCTGTGCTTAATAAGAGAAATTAT 2233
Qy      2100 GAAATPAAAAAAAAAAAA 2116
Db      2234 GAAATPAAAAAAAAAAAA 2250

RESULT 4
BC027484
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2283)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schnee,T.E., Brownstein,M.J., Udell,T.B., Toshitsuyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loqueillano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulik,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kelleman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,X.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
Schmerch,A., Schein,J.B., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
MEDLINE
PUBMED
22388257
1247932
2 (bases 1 to 2283)
Strausberg,R.
Direct Submission
Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
DNA Sequencing by: Baylor College of Medicine Human Genome

```

Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louie, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILM at: <http://image.llnl.gov>
 Series: IRAX Plate: 52 Row: f Column: 11
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES

source

1..2283
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:34809 IMAGE:5104452"
 /issue_type="Cervix, carcinoma"
 /clone_1ib="NH MGC_12"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 1..2283
 /gene="C20orf97"
 /note="Synonyms: d11103g7.3, SKIP3, TRB3, SINK"
 /db_xref="LocustID:57761"
 295..1371
 /codon_start=1
 /product="chromosome 20 open reading frame 97"
 /protein_id="AAH27484.1"
 /db_xref="GI:20071611"
 /translation="MKATPLAAPAGSLSRKKRLIEDNDLTERPVQKARSGPQRLP
 PCLLPLSEPTADRAATAVATASRLGPLYVLEPEGRAARALHCTGTEYTKVAVQ
 EALVLEPYARLPKHVAPTRPVLAGTLLAFRTGDMGMSVSRHLPPEEAA
 VLFROMKTLAHCHQGLVIRDLKCRFPADRRKIVLKNLEMSCVLTGDSLRMD
 KACRAVYSEPLISLASISGRADVWSIGVALFTMLAGHYFQSESPVLTLDGKIRG
 AALPAGLSAPARCIVRLRREPARPRTATGILHPLRQDPMPLPRLSHLMEAAQ
 VVPDGLDEAEBEEDREVVLYG"
 532..1239
 /note="S.TKc: Region: Serine/Threonine protein kinases,
 catalytic domain. Phosphotransferases of the serine or
 threonine-specific kinase subfamily. The enzymatic
 activity of these protein kinases is controlled by
 phosphorylation of specific residues in the activation
 segment of the catalytic domain, sometimes combined with
 reversible conformational changes in the C-terminal
 autoregulatory tail"
 /db_xref="CDD:c00180"

CDS

gene

misc_feature

ORIGIN

Query Match 99.0%; Score 2093.8; DB 9; Length 2283;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 2109; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

1 GGAAGCTGGCTCCGCGCGCTGCTGTAAGATCCCGGCGAGGGCTGAGCTGGGGCTGGGA 60
 134 GGAAGCGGGCTCCGCGCGCTGCTGTAAGATCCCGGCGAGGGCTGAGCTGGGGCTGGGA 193
 61 TCCCGAGCTCGGAGAGAGGCA-CGGGCGGGCCCACTGCTGCTGAGCTGGAGGCTCTGA 119
 194 TCCCGAGCTCGGAGAGAGGCA-CGGGCGGGCCCACTGCTGCTGAGCTGGAGGCTCTGA 253
 254 GCCCGCGGGCGCGCGGGCCCAAGCGGGCGAGATGCGAGCCACCCCTCTGG 313
 120 GCCCGCGGGCGCGGGCCCAAGCGGGCGAGATGCGAGCCACCCCTCTGG 179
 180 CTGCTCTCTGGGGTTCCTCTGCAAGAAAGGGTTGGAGTTGGATGACAACTTAGATA 239
 314 CTGCTCTCTGGGGTTCCTCTGCAAGAAAGGGTTGGAGTTGGATGACAACTTAGATA 373

240 CCGAGCGTCCCGTCCAGAAAGAGCTCGAGTGGGGCCCCAGCCAGACTGCCCCCTGCC 299
 374 CCGAGCGTCCCGTCCAGAAAGAGCTCGAGTGGGGCCCCAGCCAGACTGCCCCCTGCC 433
 300 TGTGGCCCTGAGGCCACCTTACTGCTCGAGATGTCGCAACTGCTGTGACCTGCTGCC 359
 434 TGTGGCCCTGAGGCCACCTTACTGCTCGAGATGTCGCAACTGCTGTGACCTGCTGCC 493
 360 GTCTTGGGCCCTTATGCTCTCTGAGAGCCGAGAGAGGGGGGGGCTTACAGAGCCCTGCC 419
 494 GTCTTGGGCCCTTATGCTCTCTGAGAGCCGAGAGAGGGGGGGGCTTACAGAGCCCTGCC 553
 420 ACTGCCCTACAGGCACTAGTATACCTGCAAGTGTACCCCTCCAGAGAGCCCTGAGCGG 479
 554 ACTGCCCTACAGGCACTAGTATACCTGCAAGTGTACCCCTCCAGAGAGCCCTGAGCGG 613
 480 TGTGAGAGCCCTTACGCGGGGCTGCCCCGACAGCATGTGCTTGCCCCACTGAGGTCC 539
 614 TGTGAGAGCCCTTACGCGGGGCTGCCCCGACAGCATGTGCTTGCCCCACTGAGGTCC 673
 540 TGGCTGTACCCAGGCTCCCTTACGCTTTTTCACCTGGAGCCATGAGGACATGACAGCC 599
 674 TGGCTGTACCCAGGCTCCCTTACGCTTTTTCACCTGGAGCCATGAGGACATGACAGCC 733
 600 TGTGCGAAACGCGCCACCGTATCCCTGAGCCTGAGGCTGCGTCTCTTCCGCAAGATGG 659
 734 TGTGCGAAACGCGCCACCGTATCCCTGAGCCTGAGGCTGCGTCTCTTCCGCAAGATGG 793
 660 CCACCGCCCTGGCGCAGTGTACCGACACGGTGTGCTGCTGCTGATCTCAAGCTGTGTC 719
 794 CCACCGCCCTGGCGCAGTGTACCGACACGGTGTGCTGCTGCTGATCTCAAGCTGTGTC 853
 720 GCTTGTCTTGGCTGACCGGTGAGAGAGAGCTGGTGTGAGAACTGAGAGACTCCT 779
 854 GCTTGTCTTGGCTGACCGGTGAGAGAGAGCTGGTGTGAGAACTGAGAGACTCCT 913
 780 GCGTCTGACTGGGGCGAGATGATCCCTGTGGAGAACAGCAGCGGTGCCAGCTTACGTGG 839
 914 GCGTCTGACTGGGGCGAGATGATCCCTGTGGAGAACAGCAGCGGTGCCAGCTTACGTGG 973
 840 GACCTGAGATATCTAGACTACGGGGCTCATCTCGGGCAAGGCAAGCCGATGTCTGAGGCC 899
 974 GACCTGAGATATCTAGACTACGGGGCTCATCTCGGGCAAGGCAAGCCGATGTCTGAGGCC 1033
 900 TGGGGGTGGCGCTCTTACCATGCTGAGCGGGCACTACCCCTTCAGAGACTCGAGACTCG 959
 1034 TGGGGGTGGCGCTCTTACCATGCTGAGCGGGCACTACCCCTTCAGAGACTCGAGACTCG 1093
 960 TCTGCTCTTGGCAAGATCGCGGGGGCTTACGCTTGGCTGCAAGGCTCTTGGGCC 1019
 1094 TCTGCTCTTGGCAAGATCGCGGGGGCTTACGCTTGGCTGCAAGGCTCTTGGGCC 1153
 1020 CTGCCCCGCTGTGCTGCTGCTGCTCTCTCTGAGAGCCAGCTGATAGCGCTACAGGCCA 1079
 1154 CTGCCCCGCTGTGCTGCTGCTGCTCTCTCTGAGAGCCAGCTGATAGCGCTACAGGCCA 1213
 1080 CAGGCACTCTCTGCAACCCCTGCTGCAAGAGCCGATGCTTACGCCCAACCCGAT 1139
 1214 CAGGCACTCTCTGCAACCCCTGCTGCAAGAGCCGATGCTTACGCCCAACCCGAT 1273
 1140 CCGATCTTGGAGAGCTGCGCAGGTGCTCTGTATGATCTGAGGCTGAGAGAGCCAGGG 1199
 1274 CCGATCTTGGAGAGCTGCGCAGGTGCTCTGTATGATCTGAGGCTGAGAGAGCCAGGG 1333
 1200 AAGAGAGGAGACAGAGAGGTGCTCTGTATGATCTGAGGCTGAGAGAGCCAGGG 1259
 1334 AAGAGAGGAGACAGAGAGGTGCTCTGTATGATCTGAGGCTGAGAGAGCCAGGG 1393
 1260 CTGCAAGAGTGAATTGAGTTGGGGGTAGCTCAAGCTTCTCTGCTCTGAACTGAG 1319
 1394 CTGCAAGAGTGAATTGAGTTGGGGGTAGCTCAAGCTTCTCTGCTCTGAACTGAG 1453
 1320 CCAAGCTTCAAGTCTTCCAGAGAGAGAGAGAGAGCTGCTGTGAGTGTGCTGTG 1379

Db 1100 GCTTTGCTTCGCTGACCTGTGAGAGAGAAAGCTGCTGTGAGAACTCTGAGAGACTCT 1159
 QY 780 GCGTGTGATCTGGGCGGAGATGATTCCTGTGGGAGCAAGCAAGCTGCTGAGCTTACGTGG 839
 Db 1160 GCGTGTGATCTGGGCGGAGATGATTCCTGTGGGAGCAAGCAAGCTGCTGAGCTTACGTGG 1219
 QY 840 GACCTGAGTACTAGCTGACGAGGCTCATATCTGGGCAAGGAGCCGATGTCTGAGCC 899
 Db 1220 GACCTGAGTACTAGCTGACGAGGCTCATATCTGGGCAAGGAGCCGATGTCTGAGCC 1279
 QY 900 TGGGCGTGGGCTTCTTACCATGCTGGCGGCACTACCCCTTTCAGAGACTGAGAGCTTG 959
 Db 1280 TGGGCGTGGGCTTCTTACCATGCTGGCGGCACTACCCCTTTCAGAGACTGAGAGCTTG 1339
 QY 960 TCCGTCTTGGGCAAGTCCGCGGGGGCTACGCTGCTGAGAGCTCTGAGCCCTCTGGGCC 1019
 Db 1340 TCTGTCTTGGGCAAGTCCGCGGGGGCTACGCTGCTGAGAGCTCTGAGCCCTCTGGGCC 1399
 QY 1020 CTGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1079
 Db 1400 CTGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1459
 QY 1080 CAGGCACTCTTCTGCAACCTCTGCTGCAAGCAAGCTGCTGCTGCTGCTGCTGCTGCTG 1139
 Db 1460 CAGGCACTCTTCTGCAACCTCTGCTGCAAGCAAGCTGCTGCTGCTGCTGCTGCTGCTG 1519
 QY 1140 CCATCTCTGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1199
 Db 1520 CCATCTCTGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1579
 QY 1200 AAGAGAGGAGAGACAGAGAGAGTGTCTGTATGCTAGAGCAACCTTACTACAGCTCAG 1259
 Db 1580 AAGAGAGGAGAGACAGAGAGAGTGTCTGTATGCTAGAGCAACCTTACTACAGCTCAG 1639
 QY 1260 CTGCGCAAGTGTGATGATTTGGGGGTAGCTGCAAGCTTCTGCTGCTGCTGCTGCTGCTG 1319
 Db 1640 CTGCGCAAGTGTGATGATTTGGGGGTAGCTGCAAGCTTCTGCTGCTGCTGCTGCTGCTG 1699
 QY 1320 CCAACCTTCTGAGCTTCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1379
 Db 1700 CCAACCTTCTGAGCTTCTGAG 1759
 QY 1380 TACACATCTGCTTGTGCTGCAACATGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1439
 Db 1760 TACACATCTGCTTGTGCTGCAACATGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1819
 QY 1440 GCGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1499
 Db 1820 GCGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1879
 QY 1500 AGAGATGACAAACTGGGATCTTGTGAGTGAACAACCTTTCATGACCATAGGTCATGCTGT 1559
 Db 1880 AGAGATGACAAACTGGGATCTTGTGAGTGAACAACCTTTCATGACCATAGGTCATGCTGT 1939
 QY 1560 CTACACTGGGATCACTTGTGACAGTGTGCGCTTCACTGATGCTGCTGCTGCTGCTGCTG 1619
 Db 1940 CTACACTGGGATCACTTGTGACAGTGTGCGCTTCACTGATGCTGCTGCTGCTGCTGCTG 1999
 QY 1620 CTGTCTCAAGGACATCTTGTGCAAAACCAACAGCTGCTTGTGATCTGATACCTTTTC 1679
 Db 2000 CTGTCTCAAGGACATCTTGTGCAAAACCAACAGCTGCTTGTGATCTGATACCTTTTC 2059
 QY 1680 AGAGAAAGGAGGATCTGCTGCAAAAGGCTCCAGGCTCTGCTGCTGCTGCTGCTGCTG 1739
 Db 2060 AGAGAAAGGAGGATCTGCTGCAAAAGGCTCCAGGCTCTGCTGCTGCTGCTGCTGCTG 2119
 QY 1740 CAAGCCAGCTCACTGTGGAACTGTGCTCCAGCACTTCTGCTGCTGATTAAGAGAT 1799
 Db 2120 CAAGCCAGCTCACTGTGGAACTGTGCTCCAGCACTTCTGCTGCTGATTAAGAGAT 2179
 QY 1800 TCTCTCTTCAAGGCTTGAAGCTTGGGCTGAGAGATGAGATCAAACTATGAGGCT 1859
 Db 2180 TCTCTCTTCAAGGCTTGAAGCTTGGGCTGAGAGATGAGATCAAACTATGAGGCT 2239

QY 1860 AGTCTTGTCTACTCAAGACTGTCTGTGAGATGAGGCTCCAGGCTCTGTCAACCATGAGGCT 1919
 Db 2240 AGTCTTGTCTACTCAAGACTGTCTGTGAGATGAGGCTCCAGGCTCTGTCAACCATGAGGCT 2299
 QY 1920 TTCTGACCTGAGACCAAGGTTGAGGAGCAGAGATTAGGACAGGCTCTGCTGAGCCACC 1979
 Db 2300 TTCTGACCTGAGACCAAGGTTGAGGAGCAGAGATTAGGACAGGCTCTGCTGAGCCACC 2359
 QY 1980 TGGAAAGTCCAGGTGGACTCTTCTGGGAGACCTTGGGGTCCAAATCCAGGTCATA 2039
 Db 2360 TGGAAAGTCCAGGTGGACTCTTCTGGGAGACCTTGGGGTCCAAATCCAGGTCATA 2419
 QY 2040 CTCTAGTTTGGATACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2099
 Db 2420 CTCTAGTTTGGATACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2479
 QY 2100 GAAATTA 2106
 Db 2480 GAAATTA 2486

RESULT 6
 BC019363
 LOCUS
 DEFINITION Homo sapiens chromosome 20 open reading frame 97, mRNA (cDNA clone
 ACCESSION BC019363
 VERSION BC019363.1 GI:17939414
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 1 (bases 1 to 2121)
 Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Burow, K.H., Scheffer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E.,
 Schermer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences

JOURNAL MEDLINE
 PUBMED
 22388257
 12477932
 2 (bases 1 to 2121)
 Strausberg, R.
 Direct Submission
 TITLE
 JOURNAL

REFERENCE
 1 (bases 1 to 2121)
 Strausberg, R.
 Direct Submission
 TITLE
 JOURNAL

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada

info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Diane Smalins, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>
Series: IRAL Plate: 8 Row: 9 Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 11056039.

FEATURES

source

1..2121
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:854 IMAGE:3528490"
/issue_type="Muscle, rhabdomyosarcoma"
/clone_1b="NIH MGC 17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..2121
/gene="C20orf97"
/note="synonyms: dJ1103g7.3, SKIP3, TRB3, SINK"
/db_xref="LocustID:57761"
108..1184
/codon_start=1
/product="chromosome 20 open reading frame 97"
/note="synonyms: dJ1103g7.3, SKIP3, TRB3, SINK"
/db_xref="GI:17939415"
/protein_id="AAH19363.1"
/db_xref="GI:17939415"
/db_xref="LocustID:57761"
/translation="MRAPPLAPAGSLSHKRRLDLNDLTERPVOKRARGSPRLP
PCLPLSPPTADRAATAVAATSRIGPVYLLRBEGRARQALHCPRTYTKYFVQ
EALVIEPARLPFHGVARPTFVLAIGTOLVAFRTIGDMSLVSRHLPBEPAA
VLRQWATLALAHQHOGLVRLKLCRFVADREKRLVLEMLDS CVLTGDDSLMD
KACAPAYGPEILISRAYSIGKADWVSLGVALFTMLAGHYFPDSEPVLLGKLRG
AYALPAGLSAPARCIVRCILREPARRLTATGILLHPWLRQDPMPLAPRSLWLEAAQ
VYVDGLDEAREEBSGDRVLYYG"

CDs

misc_feature
345..1052
/note="S_TKc: Region: Serine/Threonine protein kinases,
catalytic domain"
/db_xref="CDD:smat00220"

ORIGIN

Query Match 96.7%; Score 2046.2; DB 9; Length 2121;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2059; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 55 CTGGGATCCCGAGCTGGCAGCAGCGCA-CGGGCGGGCCCACTGTGTGCTCTGGAGG 113
DB 1 CTGGGATCCCGAGCTGGCAGCAGCGCA-CGGGCGGGCCCACTGTGTGCTCTGGAGG 60
QY 114 CTCTGAGCCCGCGCGCGCCCGGCGCCACGCGGAAAGCGCGGAGATCGAGCCACC 173
DB 61 CTCTGAGCCCGCGCGCGCCCGGCGCCACGCGGAAAGCGCGGAGATCGAGCCACC 120
QY 174 CTCTGAGCTCTCTGGGGGTTCCCTGTCCAGAGAAAGCGGTGGAGTTGATGAAACT 233
DB 121 CTCTGAGCTCTCTGGGGGTTCCCTGTCCAGAGAAAGCGGTGGAGTTGATGAAACT 180
QY 234 TAGATACGAGGTCCTCGTCAGAAAGAGAGTGAAGTGGGCCCGAGCCAGACTGCC 293
DB 181 TAGATACGAGGTCCTCGTCAGAAAGAGAGTGAAGTGGGCCCGAGCCAGACTGCC 240
QY 294 CCGTCTGTTGCTTGGAGCCCACTGCTCTCAAGATGTGCACTGCTGTGGCACTG 353
DB 241 CCGTCTGTTGCTTGGAGCCCACTGCTCTCAAGATGTGCACTGCTGTGGCACTG 300
QY 354 CCGCGGCTTGGGCGCTATGTCTCTCTGAGCCGAGAGAGGCGGGCGGCTTACCAAG 413

DB 301 CTTCCGCTTGGGCGCTATGTCTCTCTGAGCCGAGAGAGGCGGGCGGCTTACCAAG 360
QY 414 CCTGCACTGACCTTACAGGACATGATTAAGTGAAGGTATACCCGTCAGAGAGCC 473
DB 361 CCTGCACTGACCTTACAGGACATGATTAAGTGAAGGTATACCCGTCAGAGAGCC 420
QY 474 TGGCGGTGTGAGAGCCCTTACGCGCGGCTGCCCCCGACAGCATGTGCTGGGCCACTG 533
DB 421 TGGCGGTGTGAGAGCCCTTACGCGCGGCTGCCCCCGACAGCATGTGCTGGGCCACTG 480
QY 534 AGTCTCTGCTGATACCAAGCTCTCTAGCCTTTTTCATGAGCCCATGAGGCTGCTTCCGCG 593
DB 481 AGTCTCTGCTGATACCAAGCTCTCTAGCCTTTTTCATGAGCCCATGAGGCTGCTTCCGCG 540
QY 594 ACAGCTGTGAGAGCGCCACCGCATCTGAGCGCTGAGGCTGCGGCTCTTCCGCG 653
DB 541 ACAGCTGTGAGAGCGCCACCGCATCTGAGCGCTGAGGCTGCGGCTCTTCCGCG 600
QY 654 AGATGGCCACCGCCTTGGCGCACTGTACACAGCAAGCTGTGCTCTGCTGATCTCAAGC 713
DB 601 AGATGGCCACCGCCTTGGCGCACTGTACACAGCAAGCTGTGCTCTGCTGATCTCAAGC 660
QY 714 TGTGTGCTTGTCTTCTGCTGACCGTGAAGAGAGAGTGTGCTGGAACCTTGAGG 773
DB 661 TGTGTGCTTGTCTTCTGCTGACCGTGAAGAGAGTGTGCTGGAACCTTGAGG 720
QY 774 ACTCTGCTGTGCTGACCTGGGCGAGATGATTCCTTGGGACAAAGCGGTGCCAGCCT 833
DB 721 ACTCTGCTGTGCTGACCTGGGCGAGATGATTCCTTGGGACAAAGCGGTGCCAGCCT 780
QY 834 ACTGGGACCTGAGATTAAGTCAAGCTCAAGGCGCTCATTAAGTGGGAGAGCGGATGTCT 893
DB 781 ACTGGGACCTGAGATTAAGTCAAGCTCAAGGCGCTCATTAAGTGGGAGAGCGGATGTCT 840
QY 894 GGAAGCTGGGCGTGGCGCTTACAGATGCTGGCGGCGCAATACCCCTTCCAGAGTCTCG 953
DB 841 GGAAGCTGGGCGTGGCGCTTACAGATGCTGGCGGCGCAATACCCCTTCCAGAGTCTCG 900
QY 954 AGCTGTCTGTCTTCTGGCAAGATCCGCGGGGCGCTACAGCTTGTCTGAGGCTCT 1013
DB 901 AGCTGTCTGTCTTCTGGCAAGATCCGCGGGGCGCTACAGCTTGTCTGAGGCTCT 960
QY 1014 CGGCGCTGCGCGCTGTGTGCTGTGCTGCTCTCTTGTGTGAGAGCCAGTGAACGCTCA 1073
DB 961 CGGCGCTGCGCGCTGTGTGCTGTGCTGCTCTCTTGTGTGAGAGCCAGTGAACGCTCA 1020
QY 1074 CAGCCAGGAGATCTCTCTGCAACCCCTGGCTGCGACAGGACCGGATGCCCTTACGCCCA 1133
DB 1021 CAGCCAGGAGATCTCTCTGCAACCCCTGGCTGCGACAGGACCGGATGCCCTTACGCCCA 1080
QY 1134 CCGGATCCCATCTTGGAGGCTGCGCAGTGGTCCCTGATGAGACTGGGCTGAGCAAG 1193
DB 1081 CCGGATCCCATCTTGGAGGCTGCGCAGTGGTCCCTGATGAGACTGGGCTGAGCAAG 1140
QY 1194 CCAAGGAGAGAGAGAGAGACAGAGAAATGTTCTGTATGCTGAGCAACCTTAACAC 1253
DB 1141 CCAAGGAGAGAGAGAGAGACAGAGAAATGTTCTGTATGCTGAGCAACCTTAACAC 1200
QY 1254 GCTCAGCTGCGCAAGTGAATGAGTTGGGGGTGAGCTCAAGCCCTTCTCGGCTGTA 1313
DB 1201 GCTCAGCTGCGCAAGTGAATGAGTTGGGGGTGAGCTCAAGCCCTTCTCGGCTGTA 1260
QY 1314 ACTGAGCCAAACCTTCAGTCTTCCAGAGAGAGAGAGAGAGCCCTGTGTGAGTGT 1373
DB 1261 ACTGAGCCAAACCTTCAGTCTTCCAGAGAGAGAGAGAGAGCCCTGTGTGAGTGT 1320
QY 1374 GCTGTGTACATCTGCTTTGTTCCACACATGACATGCTCTGCTGGTGTATCAAG 1433
DB 1321 GCTGTGTACATCTGCTTTGTTCCACACATGACATGCTCTGCTGGTGTATCAAG 1380
QY 1434 TGGCAAGCCCTGTTCTCGGTGTGAGGAGTGAAGAGTGAAGAGCAATATCTCCCT 1493


```

Db      1381 TGGCAAGCCCTGTTCTCGGTGCTGGAGATACAGCATGAGCAAGACATATTCCT 1440
QY      1494 GCTCACAGAGATACAAACCTGGATCTTTAGCTGACAACTTTTCCATGACCAATAGGT 1553
Db      1441 GCTCACAGAGATACAAACCTGGATCTTTAGCTGACAACTTTTCCATGACCAATAGGT 1500
QY      1554 CACTGTCTACACAGGATACACTTTGTACAGTGTGGCTCCCACTGATGTGGTGCTCAG 1613
Db      1501 CACTGTCTACACAGGATACACTTTGTACAGTGTGGCTCCCACTGATGTGGTGCTCAG 1560
QY      1614 GCACTCTGTCCACAGACAAATCCCTTTCACAAACAAACGAGCTCTTGTATCTTTATC 1673
Db      1561 GCACCTCTGTCCACAGACAAATCCCTTTCACAAACAAACGAGCTCTTGTATCTTTATC 1620
QY      1674 CTTTTCAGAGAAAGGAGGATACCTGTGCAAGGCTCCAGGCTCTCCCTGCAATC 1733
Db      1621 CTTTTCAGAGAAAGGAGGATACCTGTGCAAGGCTCCAGGCTCTCCCTGCAATC 1680
QY      1734 AGGACCCAGCCAGCTCACTGTGGAACTGTGTCCAGCATCTGTCTCTGATTA 1793
Db      1681 AGGACCCAGCCAGCTCACTGTGGAACTGTGTCCAGCATCTGTCTCTGATTA 1740
QY      1794 AGGATTTCTCTTCCAGGCTTAAGCTGGGATTGGGCCAGAGATTAAGATCCAACTAT 1853
Db      1741 AGGATTTCTCTTCCAGGCTTAAGCTGGGATTGGGCCAGAGATTAAGATCCAACTAT 1800
QY      1854 GAGGCTAGTTCTGTCTTAATCAAGACTGTCTGGAATGAGGCTCCAGGCTCTCAACA 1913
Db      1801 GAGGCTAGTTCTGTCTTAATCAAGACTGTCTGGAATGAGGCTCTCAACA 1860
QY      1914 TGGGGCTTCTGACCTGACGACCAAGGTTGAGGACAGAGATTAGGAGGCTCTCTGTTG 1973
Db      1861 TGGGGCTTCTGACCTGACGACCAAGGTTGAGGACAGAGATTAGGAGGCTCTCTGTTG 1920
QY      1974 GCCACCTCGAAAGTCCAGGTGGGACTCTTTCGGGACACTTGGGTTCCAAATCCAGG 2033
Db      1921 GCCACCTCGAAAGTCCAGGTGGGACTCTTTCGGGACACTTGGGTTCCAAATCCAGG 1980
QY      2034 TCATATCTAGGTTTGGATACATGAGATGATGTTACTGTGGCCATTAAGAGGAG 2093
Db      1981 TCATATCTAGGTTTGGATACATGAGATGATGTTACTGTGGCCATTAAGAGGAG 2040
QY      2094 AATTATGAATATATATATATATATATATATATATATATATATATATATATATATAT 2116
Db      2041 AATTATGAATATATATATATATATATATATATATATATATATATATATATATAT 2063

RESULT 7
LOCUS   HS1103G7              153170 bp    DNA             linear    PRI 08-FEB-2001
DEFINITION Human DNA sequence from clone RPS-1103G7 on chromosome 20p12.2-13.
           Contains up to three novel genes, the gene for a novel protein
           similar to mouse VMP, the gene for a novel protein kinase domain
           containing protein similar to phosphoprotein C8W and rat N1PK, and
           the SOX2 gene for SRY (sex-determining region Y)-box 22. Contains
           five CpG islands, ESTs, STSs and GSSs, complete sequence.
ACCESSION AL034548.25 GI:7263904
VERSION   AL034548
KEYWORDS  HTG, CpG island, N1PK, protein kinase, SOX2, SRY, VMP.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 153170)
AUTHORS   Blakey,S.
TITLE     Direct Submission
JOURNAL   Submitted (05-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
           CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
COMMENT   On Mar 19, 2000 this sequence version replaced gi:5541861.
           During sequence assembly data is compared from overlapping clones.
           Where differences are found these are annotated as variations
           together with a note of the overlapping clone name. Note that the

```

```

FEATURES
source
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Infomation
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep This sequence is
the entire, insert of clone RPS-1103G7 This sequence was finished as
follows unless otherwise noted: all regions were either
double-stranded or sequenced with an alternate chemistry or covered
by high quality data (i.e., phred quality >= 30); an attempt was
made to resolve all sequencing problems, such as compressions and
repeats: all regions were covered by at least one plasmid subclone
or more than one M13 subclone, and the assembly was confirmed by
restriction digest. This sequence was generated from part of
bacterial clone contigs of human chromosome 20, constructed by the
Sanger Centre Chromosome 20 Mapping Group. Further information can
be found at http://www.sanger.ac.uk/HGP/Chr20
RPS-1103G7 is from the library RPCI-5 constructed by the group of
Piet de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.
location/Qualifiers
1..153170
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="p12.2-13"
/clone="RPS-1103G7"
/clone_1fb="RPCI-5"
1..61
/feature="Alu repeat: matches 2..62 of consensus"
79..129
/feature="L1M4 repeat: matches 6221..6271 of consensus"
130..173
/feature="L1M3 repeat: matches 7715..7921 of consensus"
175..873
/feature="L1M5 repeat: matches 5519..6217 of consensus"
902..1109
/feature="L1M2 repeat: matches 115..298 of consensus"
1110..1480
/feature="L1M1 repeat: matches 115..298 of consensus"
1502..1733
/feature="L1M4 repeat: matches 115..298 of consensus"
1734..1888
/feature="L1M3 repeat: matches 115..298 of consensus"
1919..2101
/feature="L1M2 repeat: matches 115..298 of consensus"
2154..2193
/feature="L1M1 repeat: matches 115..298 of consensus"
2194..2243
/feature="L1M4 repeat: matches 115..298 of consensus"
2244..2293
/feature="L1M3 repeat: matches 115..298 of consensus"
2294..2343
/feature="L1M2 repeat: matches 115..298 of consensus"
2344..2393
/feature="L1M1 repeat: matches 115..298 of consensus"
2394..2443
/feature="L1M4 repeat: matches 115..298 of consensus"
2444..2493
/feature="L1M3 repeat: matches 115..298 of consensus"
2494..2543
/feature="L1M2 repeat: matches 115..298 of consensus"
2544..2593
/feature="L1M1 repeat: matches 115..298 of consensus"
2594..2643
/feature="L1M4 repeat: matches 115..298 of consensus"
2644..2693
/feature="L1M3 repeat: matches 115..298 of consensus"
2694..2743
/feature="L1M2 repeat: matches 115..298 of consensus"
2744..2793
/feature="L1M1 repeat: matches 115..298 of consensus"
2794..2843
/feature="L1M4 repeat: matches 115..298 of consensus"
2844..2893
/feature="L1M3 repeat: matches 115..298 of consensus"
2894..2943
/feature="L1M2 repeat: matches 115..298 of consensus"
2944..2993
/feature="L1M1 repeat: matches 115..298 of consensus"
2994..3043
/feature="L1M4 repeat: matches 115..298 of consensus"
3044..3093
/feature="L1M3 repeat: matches 115..298 of consensus"
3094..3143
/feature="L1M2 repeat: matches 115..298 of consensus"
3144..3193
/feature="L1M1 repeat: matches 115..298 of consensus"
3194..3243
/feature="L1M4 repeat: matches 115..298 of consensus"
3244..3293
/feature="L1M3 repeat: matches 115..298 of consensus"
3294..3343
/feature="L1M2 repeat: matches 115..298 of consensus"
3344..3393
/feature="L1M1 repeat: matches 115..298 of consensus"
3394..3443
/feature="L1M4 repeat: matches 115..298 of consensus"
3444..3493
/feature="L1M3 repeat: matches 115..298 of consensus"
3494..3543
/feature="L1M2 repeat: matches 115..298 of consensus"
3544..3593
/feature="L1M1 repeat: matches 115..298 of consensus"
3594..3643
/feature="L1M4 repeat: matches 115..298 of consensus"
3644..3693
/feature="L1M3 repeat: matches 115..298 of consensus"
3694..3743
/feature="L1M2 repeat: matches 115..298 of consensus"
3744..3793
/feature="L1M1 repeat: matches 115..298 of consensus"
3794..3843
/feature="L1M4 repeat: matches 115..298 of consensus"
3844..3893
/feature="L1M3 repeat: matches 115..298 of consensus"
3894..3943
/feature="L1M2 repeat: matches 115..298 of consensus"
3944..3993
/feature="L1M1 repeat: matches 115..298 of consensus"
3994..4043
/feature="L1M4 repeat: matches 115..298 of consensus"
4044..4093
/feature="L1M3 repeat: matches 115..298 of consensus"
4094..4143
/feature="L1M2 repeat: matches 115..298 of consensus"
4144..4193
/feature="L1M1 repeat: matches 115..298 of consensus"
4194..4243
/feature="L1M4 repeat: matches 115..298 of consensus"
4244..4293
/feature="L1M3 repeat: matches 115..298 of consensus"
4294..4343
/feature="L1M2 repeat: matches 115..298 of consensus"
4344..4393
/feature="L1M1 repeat: matches 115..298 of consensus"
4394..4443
/feature="L1M4 repeat: matches 115..298 of consensus"
4444..4493
/feature="L1M3 repeat: matches 115..298 of consensus"
4494..4543
/feature="L1M2 repeat: matches 115..298 of consensus"
4544..4593
/feature="L1M1 repeat: matches 115..298 of consensus"
4594..4643
/feature="L1M4 repeat: matches 115..298 of consensus"
4644..4693
/feature="L1M3 repeat: matches 115..298 of consensus"
4694..4743
/feature="L1M2 repeat: matches 115..298 of consensus"
4744..4793
/feature="L1M1 repeat: matches 115..298 of consensus"
4794..4843
/feature="L1M4 repeat: matches 115..298 of consensus"
4844..4893
/feature="L1M3 repeat: matches 115..298 of consensus"
4894..4943
/feature="L1M2 repeat: matches 115..298 of consensus"
4944..4993
/feature="L1M1 repeat: matches 115..298 of consensus"
4994..5043
/feature="L1M4 repeat: matches 115..298 of consensus"
5044..5093
/feature="L1M3 repeat: matches 115..298 of consensus"
5094..5143
/feature="L1M2 repeat: matches 115..298 of consensus"
5144..5193
/feature="L1M1 repeat: matches 115..298 of consensus"
5194..5243
/feature="L1M4 repeat: matches 115..298 of consensus"
5244..5293
/feature="L1M3 repeat: matches 115..298 of consensus"
5294..5343
/feature="L1M2 repeat: matches 115..298 of consensus"
5344..5393
/feature="L1M1 repeat: matches 115..298 of consensus"
5394..5443
/feature="L1M4 repeat: matches 115..298 of consensus"
5444..5493
/feature="L1M3 repeat: matches 115..298 of consensus"
5494..5543
/feature="L1M2 repeat: matches 115..298 of consensus"
5544..5593
/feature="L1M1 repeat: matches 115..298 of consensus"
5594..5643
/feature="L1M4 repeat: matches 115..298 of consensus"
5644..5693
/feature="L1M3 repeat: matches 115..298 of consensus"
5694..5743
/feature="L1M2 repeat: matches 115..298 of consensus"
5744..5793
/feature="L1M1 repeat: matches 115..298 of consensus"
5794..5843
/feature="L1M4 repeat: matches 115..298 of consensus"
5844..5893
/feature="L1M3 repeat: matches 115..298 of consensus"
5894..5943
/feature="L1M2 repeat: matches 115..298 of consensus"
5944..5993
/feature="L1M1 repeat: matches 115..298 of consensus"
5994..6043
/feature="L1M4 repeat: matches 115..298 of consensus"
6044..6093
/feature="L1M3 repeat: matches 115..298 of consensus"
6094..6143
/feature="L1M2 repeat: matches 115..298 of consensus"
6144..6193
/feature="L1M1 repeat: matches 115..298 of consensus"
6194..6243
/feature="L1M4 repeat: matches 115..298 of consensus"
6244..6293
/feature="L1M3 repeat: matches 115..298 of consensus"
6294..6343
/feature="L1M2 repeat: matches 115..298 of consensus"
6344..6393
/feature="L1M1 repeat: matches 115..298 of consensus"
6394..6443
/feature="L1M4 repeat: matches 115..298 of consensus"
6444..6493
/feature="L1M3 repeat: matches 115..298 of consensus"
6494..6543
/feature="L1M2 repeat: matches 115..298 of consensus"
6544..6593
/feature="L1M1 repeat: matches 115..298 of consensus"
6594..6643
/feature="L1M4 repeat: matches 115..298 of consensus"
6644..6693
/feature="L1M3 repeat: matches 115..298 of consensus"
6694..6743
/feature="L1M2 repeat: matches 115..298 of consensus"
6744..6793
/feature="L1M1 repeat: matches 115..298 of consensus"
6794..6843
/feature="L1M4 repeat: matches 115..298 of consensus"
6844..6893
/feature="L1M3 repeat: matches 115..298 of consensus"
6894..6943
/feature="L1M2 repeat: matches 115..298 of consensus"
6944..6993
/feature="L1M1 repeat: matches 115..298 of consensus"
6994..7043
/feature="L1M4 repeat: matches 115..298 of consensus"
7044..7093
/feature="L1M3 repeat: matches 115..298 of consensus"
7094..7143
/feature="L1M2 repeat: matches 115..298 of consensus"
7144..7193
/feature="L1M1 repeat: matches 115..298 of consensus"
7194..7243
/feature="L1M4 repeat: matches 115..298 of consensus"
7244..7293
/feature="L1M3 repeat: matches 115..298 of consensus"
7294..7343
/feature="L1M2 repeat: matches 115..298 of consensus"
7344..7393
/feature="L1M1 repeat: matches 115..298 of consensus"
7394..7443
/feature="L1M4 repeat: matches 115..298 of consensus"
7444..7493
/feature="L1M3 repeat: matches 115..298 of consensus"
7494..7543
/feature="L1M2 repeat: matches 115..298 of consensus"
7544..7593
/feature="L1M1 repeat: matches 115..298 of consensus"
7594..7643
/feature="L1M4 repeat: matches 115..298 of consensus"
7644..7693
/feature="L1M3 repeat: matches 115..298 of consensus"
7694..7743
/feature="L1M2 repeat: matches 115..298 of consensus"
7744..7793
/feature="L1M1 repeat: matches 115..298 of consensus"
7794..7843
/feature="L1M4 repeat: matches 115..298 of consensus"
7844..7893
/feature="L1M3 repeat: matches 115..298 of consensus"
7894..7943
/feature="L1M2 repeat: matches 115..298 of consensus"
7944..7993
/feature="L1M1 repeat: matches 115..298 of consensus"
7994..8043
/feature="L1M4 repeat: matches 115..298 of consensus"
8044..8093
/feature="L1M3 repeat: matches 115..298 of consensus"
8094..8143
/feature="L1M2 repeat: matches 115..298 of consensus"
8144..8193
/feature="L1M1 repeat: matches 115..298 of consensus"
8194..8243
/feature="L1M4 repeat: matches 115..298 of consensus"
8244..8293
/feature="L1M3 repeat: matches 115..298 of consensus"
8294..8343
/feature="L1M2 repeat: matches 115..298 of consensus"
8344..8393
/feature="L1M1 repeat: matches 115..298 of consensus"
8394..8443
/feature="L1M4 repeat: matches 115..298 of consensus"
8444..8493
/feature="L1M3 repeat: matches 115..298 of consensus"
8494..8543
/feature="L1M2 repeat: matches 115..298 of consensus"
8544..8593
/feature="L1M1 repeat: matches 115..298 of consensus"
8594..8643
/feature="L1M4 repeat: matches 115..298 of consensus"
8644..8693
/feature="L1M3 repeat: matches 115..298 of consensus"
8694..8743
/feature="L1M2 repeat: matches 115..298 of consensus"
8744..8793
/feature="L1M1 repeat: matches 115..298 of consensus"
8794..8843
/feature="L1M4 repeat: matches 115..298 of consensus"
8844..8893
/feature="L1M3 repeat: matches 115..298 of consensus"
8894..8943
/feature="L1M2 repeat: matches 115..298 of consensus"
8944..8993
/feature="L1M1 repeat: matches 115..298 of consensus"
8994..9043
/feature="L1M4 repeat: matches 115..298 of consensus"
9044..9093
/feature="L1M3 repeat: matches 115..298 of consensus"
9094..9143
/feature="L1M2 repeat: matches 115..298 of consensus"
9144..9193
/feature="L1M1 repeat: matches 115..298 of consensus"
9194..9243
/feature="L1M4 repeat: matches 115..298 of consensus"
9244..9293
/feature="L1M3 repeat: matches 115..298 of consensus"
9294..9343
/feature="L1M2 repeat: matches 115..298 of consensus"
9344..9393
/feature="L1M1 repeat: matches 115..298 of consensus"
9394..9443
/feature="L1M4 repeat: matches 115..298 of consensus"
9444..9493
/feature="L1M3 repeat: matches 115..298 of consensus"
9494..9543
/feature="L1M2 repeat: matches 115..298 of consensus"
9544..9593
/feature="L1M1 repeat: matches 115..298 of consensus"
9594..9643
/feature="L1M4 repeat: matches 115..298 of consensus"
9644..9693
/feature="L1M3 repeat: matches 115..298 of consensus"
9694..9743
/feature="L1M2 repeat: matches 115..298 of consensus"
9744..9793
/feature="L1M1 repeat: matches 115..298 of consensus"
9794..9843
/feature="L1M4 repeat: matches 115..298 of consensus"
9844..9893
/feature="L1M3 repeat: matches 115..298 of consensus"
9894..9943
/feature="L1M2 repeat: matches 115..298 of consensus"
9944..9993
/feature="L1M1 repeat: matches 115..298 of consensus"
9994..10043
/feature="L1M4 repeat: matches 115..298 of consensus"
10044..10093
/feature="L1M3 repeat: matches 115..298 of consensus"
10094..10143
/feature="L1M2 repeat: matches 115..298 of consensus"
10144..10193
/feature="L1M1 repeat: matches 115..298 of consensus"
10194..10243
/feature="L1M4 repeat: matches 115..298 of consensus"
10244..10293
/feature="L1M3 repeat: matches 115..298 of consensus"
10294..10343
/feature="L1M2 repeat: matches 115..298 of consensus"
10344..10393
/feature="L1M1 repeat: matches 115..298 of consensus"
10394..10443
/feature="L1M4 repeat: matches 115..298 of consensus"
10444..10493
/feature="L1M3 repeat: matches 115..298 of consensus"
10494..10543
/feature="L1M2 repeat: matches 115..298 of consensus"
10544..10593
/feature="L1M1 repeat: matches 115..298 of consensus"
10594..10643
/feature="L1M4 repeat: matches 115..298 of consensus"
10644..10693
/feature="L1M3 repeat: matches 115..298 of consensus"
10694..10743
/feature="L1M2 repeat: matches 115..298 of consensus"
10744..10793
/feature="L1M1 repeat: matches 115..298 of consensus"
10794..10843
/feature="L1M4 repeat: matches 115..298 of consensus"
10844..10893
/feature="L1M3 repeat: matches 115..298 of consensus"
10894..10943
/feature="L1M2 repeat: matches 115..298 of consensus"
10944..10993
/feature="L1M1 repeat: matches 115..298 of consensus"
10994..11043
/feature="L1M4 repeat: matches 115..298 of consensus"
11044..11093
/feature="L1M3 repeat: matches 115..298 of consensus"
11094..11143
/feature="L1M2 repeat: matches 115..298 of consensus"
11144..11193
/feature="L1M1 repeat: matches 115..298 of consensus"
11194..11243
/feature="L1M4 repeat: matches 115..298 of consensus"
11244..11293
/feature="L1M3 repeat: matches 115..298 of consensus"
11294..11343
/feature="L1M2 repeat: matches 115..298 of consensus"
11344..11393
/feature="L1M1 repeat: matches 115..298 of consensus"
11394..11443
/feature="L1M4 repeat: matches 115..298 of consensus"
11444..11493
/feature="L1M3 repeat: matches 115..298 of consensus"
11494..11543
/feature="L1M2 repeat: matches 115..298 of consensus"
11544..11593
/feature="L1M1 repeat: matches 115..298 of consensus"
11594..11643
/feature="L1M4 repeat: matches 115..298 of consensus"
11644..11693
/feature="L1M3 repeat: matches 115..298 of consensus"
11694..11743
/feature="L1M2 repeat: matches 115..298 of consensus"
11744..11793
/feature="L1M1 repeat: matches 115..298 of consensus"
11794..11843
/feature="L1M4 repeat: matches 115..298 of consensus"
11844..11893
/feature="L1M3 repeat: matches 115..298 of consensus"
11894..11943
/feature="L1M2 repeat: matches 115..298 of consensus"
11944..11993
/feature="L1M1 repeat: matches 115..298 of consensus"
11994..12043
/feature="L1M4 repeat: matches 115..298 of consensus"
12044..12093
/feature="L1M3 repeat: matches 115..298 of consensus"
12094..12143
/feature="L1M2 repeat: matches 115..298 of consensus"
12144..12193
/feature="L1M1 repeat: matches 115..298 of consensus"
12194..12243
/feature="L1M4 repeat: matches 115..298 of consensus"
12244..12293
/feature="L1M3 repeat: matches 115..298 of consensus"
12294..12343
/feature="L1M2 repeat: matches 115..298 of consensus"
12344..12393
/feature="L1M1 repeat: matches 115..298 of consensus"
12394..12443
/feature="L1M4 repeat: matches 115..298 of consensus"
12444..12493
/feature="L1M3 repeat: matches 115..298 of consensus"
12494..12543
/feature="L1M2 repeat: matches 115..298 of consensus"
12544..12593
/feature="L1M1 repeat: matches 115..298 of consensus"
12594..12643
/feature="L1M4 repeat: matches 115..298 of consensus"
12644..12693
/feature="L1M3 repeat: matches 115..298 of consensus"
12694..12743
/feature="L1M2 repeat: matches 115..298 of consensus"
12744..12793
/feature="L1M1 repeat: matches 115..298 of consensus"
12794..12843
/feature="L1M4 repeat: matches 115..298 of consensus"
12844..12893
/feature="L1M3 repeat: matches 115..298 of consensus"
12894..12943
/feature="L1M2 repeat: matches 115..298 of consensus"
12944..12993
/feature="L1M1 repeat: matches 115..298 of consensus"
12994..13043
/feature="L1M4 repeat: matches 115..298 of consensus"
13044..13093
/feature="L1M3 repeat: matches 115..298 of consensus"
13094..13143
/feature="L1M2 repeat: matches 115..298 of consensus"
13144..13193
/feature="L1M1 repeat: matches 115..298 of consensus"
13194..13243
/feature="L1M4 repeat: matches 115..298 of consensus"
13244..13293
/feature="L1M3 repeat: matches 115..298 of consensus"
13294..13343
/feature="L1M2 repeat: matches 115..298 of consensus"
13344..13393
/feature="L1M1 repeat: matches 115..298 of consensus"
13394..13443
/feature="L1M4 repeat: matches 115..298 of consensus"
13444..13493
/feature="L1M3 repeat: matches 115..298 of consensus"
13494..13543
/feature="L1M2 repeat: matches 115..298 of consensus"
13544..13593
/feature="L1M1 repeat: matches 115..298 of consensus"
13594..13643
/feature="L1M4 repeat: matches 115..298 of consensus"
13644..13693
/feature="L1M3 repeat: matches 115..298 of consensus"
13694..13743
/feature="L1M2 repeat: matches 115..298 of consensus"
13744..13793
/feature="L1M1 repeat: matches 115..298 of consensus"
13794..13843
/feature="L1M4 repeat: matches 115..298 of consensus"
13844..13893
/feature="L1M3 repeat: matches 115..298 of consensus"
13894..13943
/feature="L1M2 repeat: matches 115..298 of consensus"
13944..13993
/feature="L1M1 repeat: matches 115..298 of consensus"
13994..14043
/feature="L1M4 repeat: matches 115..298 of consensus"
14044..14093
/feature="L1M3 repeat: matches 115..298 of consensus"
14094..14143
/feature="L1M2 repeat: matches 115..298 of consensus"
14144..14193
/feature="L1M1 repeat: matches 115..298 of consensus"
14194..14243
/feature="L1M4 repeat: matches 115..298 of consensus"
14244..14293
/feature="L1M3 repeat: matches 115..298 of consensus"
14294..14343
/feature="L1M2 repeat: matches 115..298 of consensus"
14344..14393
/feature="L1M1 repeat: matches 115..298 of consensus"
14394..14443
/feature="L1M4 repeat: matches 115..298 of consensus"
14444..14493
/feature="L1M3 repeat: matches 115..298 of consensus"
14494..14543
/feature="L1M2 repeat: matches 115..298 of consensus"
14544..14593
/feature="L1M1 repeat: matches 115..298 of consensus"
14594..14643
/feature="L1M4 repeat: matches 115..298 of consensus"
14644..14693
/feature="L1M3 repeat: matches 115..298 of consensus"
14694..14743
/feature="L1M2 repeat: matches 115..298 of consensus"
14744..14793
/feature="L1M1 repeat: matches 115..298 of consensus"
14794..14843
/feature="L1M4 repeat: matches 115..298 of consensus"
14844..14893
/feature="L1M3 repeat: matches 115..298 of consensus"
14894..14943
/feature="L1M2 repeat: matches 115..298 of consensus"
14944..14993
/feature="L1M1 repeat: matches 115..298 of consensus"
14994..15043
/feature="L1M4 repeat: matches 115..298 of consensus"
15044..15093
/feature="L1M3 repeat: matches 115..298 of consensus"
15094..15143
/feature="L1M2 repeat: matches 115..298 of consensus"
15144..15193
/feature="L1M1 repeat: matches 115..298 of consensus"
15194..15243
/feature="L1M4 repeat: matches 115..298 of consensus"
15244..15293
/feature="L1M3 repeat: matches 115..298 of consensus"
15294..15343
/feature="L1M2 repeat: matches 115..298 of consensus"
15344..15393
/feature="L1M1 repeat: matches 115..298 of consensus"
15394..15443
/feature="L1M4 repeat: matches 115..298 of consensus"
15444..15493
/feature="L1M3 repeat: matches 115..298 of consensus"
15494..15543
/feature="L1M2 repeat: matches 115..298 of consensus"
15544..15593
/feature="L1M1 repeat: matches 115..298 of consensus"
15594..15643
/feature="L1M4 repeat: matches 115..298 of consensus"
15644..15693
/feature="L1M3 repeat: matches 115..298 of consensus"
15694..15743
/feature="L1M2 repeat: matches 115..298 of consensus"
15744..15793
/feature="L1M1 repeat: matches 115..298 of consensus"
15794..15843
/feature="L1M4 repeat: matches 115..298 of consensus"
15844..15893
/feature="L1M3 repeat: matches 115..298 of consensus"
15894..15943
/feature="L1M2 repeat: matches 115..298 of consensus"
15944..15993
/feature="L1M1 repeat: matches 115..298 of consensus"
15994..16043
/feature="L1M4 repeat: matches 115..298 of consensus"
16044..16093
/feature="L1M3 repeat: matches 115..298 of consensus"
16094..16143
/feature="L1M2 repeat: matches 115..298 of consensus"
16144..16193
/feature="L1M1 repeat: matches 115..298 of consensus"
16194..16243
/feature="L1M4 repeat: matches 115..298 of consensus"
16244..16293
/feature="L1M3 repeat: matches 115..298 of consensus"
16294..16343
/feature="L1M2 repeat: matches 115..298 of consensus"
16344..16393
/feature="L1M1 repeat: matches 115..298 of consensus"
16394..16443
/feature="L1M4 repeat: matches 115..298 of consensus"
16444..16493
/feature="L1M3 repeat: matches 115..298 of consensus"
16494..16543
/feature="L1M2 repeat: matches 115..298 of consensus"
16544..16593
/feature="L1M1 repeat: matches 115..298 of consensus"
16594..16643
/feature="L1M4 repeat: matches 115..298 of consensus"
16644..16693
/feature="L1M3 repeat: matches 115..298 of consensus"
16694..16743
/feature="L1M2 repeat: matches 115..298 of consensus"
16744..16793
/feature="L1M1 repeat: matches 115..298 of consensus"
16794..16843
/feature="L1M4 repeat: matches 115..298 of consensus"
16844..16893
/feature="L1M3 repeat: matches 115..298 of consensus"
16894..16943
/feature="L1M2 repeat: matches 115..298 of consensus"
16944..16993
/feature="L1M1 repeat: matches 115..298 of consensus"
16994..17043
/feature="L1M4 repeat: matches 115..298 of consensus"
17044..17093
/feature="L1M3 repeat: matches 115..298 of consensus"
17094..17143
/feature="L1M2 repeat: matches 115..298 of consensus"
17144..17193
/feature="L1M1 repeat: matches 115..298 of consensus"
17194..17243
/feature="L1M4 repeat: matches 115..298 of consensus"
17244..17293
/feature="L1M3 repeat: matches 115..298 of consensus"
17294..17343
/feature="L1M2 repeat: matches 115..298 of consensus"
17344..17393
/feature="L1M1 repeat: matches 115..298 of consensus"
17394..17443
/feature="L1M4 repeat: matches 115..298
```


QY 1061 GCTGAACGGCTCAGAGCCAGGGATCTCTCTGACACCCCTGGCTGGGACAGACCCGATG 1120
Db 901 GCTGAACGGCTCAGAGCCAGGGATCTCTCTGACACCCCTGGCTGGGACAGACCCGATG 960
QY 1121 CCCTTAGCCCAACCCGATCCATCTCTGGAGGCTGCCAGGTTGCTCCCTGATGACTG 1180
Db 961 CCCTTAGCCCAACCCGATCCATCTCTGGAGGCTGCCAGGTTGCTCCCTGATGACTG 1020
QY 1181 GGGCTGAGAGAGCCGAGGAG 1237
Db 1021 GGGCTGAGAGAGCCGAGGAG 1077

RESULT 9
AY335754 1077 bp mRNA linear SYN 15-OCT-2003
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

AY335754 1077 bp mRNA linear SYN 15-OCT-2003
Synthetic construct Homo sapiens chromosome 20 open reading frame
AY335754
AY335754.1 GI:33304146
FLI_CDNA.
synthetic construct
synthetic construct
artificial sequences.
1 (bases 1 to 1077)
Park,J., Rolfs,A., Hu,Y., Shen,B., Vannberg,F., Moreira,D.,
Kelley,T., Zuo,D., Raphael,J., Baqui,M., Jepson,D., Harlow,E.,
Labber,J., and Brizuela,L.
Cloning of human full-length CDS FLEXGene kinases in
recombinational vector system
Unpublished
2 (bases 1 to 1077)
Park,J., Rolfs,A., Hu,Y., Shen,B., Vannberg,F., Moreira,D.,
Kelley,T., Zuo,D., Raphael,J., Baqui,M., Jepson,D., Harlow,E.,
Labber,J., and Brizuela,L.
Direct Submission
Submitted (02-JUL-2003) Biological Chemistry and Molecular
Pharmacology, Harvard Institute of Proteomics, 320 Charles St.,
Cambridge, MA 02141-2023, USA
This CDS clone is a part of a collection of human full-length
expression clones generated by Harvard Institute of Proteomics.
Each CDS has been cloned without stop-codon (to allow fusion with
C-terminal tag). The CDS has been directionally cloned using BD
In-Fusion(TM) cloning system between the SalI and HindIII sites of
the pDNR-Dual vector. Additional sequences in the clone: 'ACC'
after SalI site and before 'ATG' to provide Kozak consensus
sequence; 'GG' after last codon and before HindIII site to maintain
reading frame.
Location/Qualifiers
1..1077
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="FLH000134.01L"
/clone_1id="MGC clone templates"
/lab_host="DH5alpha TI resistant"
/note="Vector: pDNR-Dual"
1..>1077
/gene="C20orf97"
1..>1077
/gene="C20orf97"
/note="Mutations: Stop->Leu"
/transl_table=1
/product="chromosome 20 open reading frame 97"
/protein_id="AA02581.1"
/db_xref="GI:33304147"
/translation="MRATPLAAGSLSRKRLDLNDLTERPVQKARSQPPRP
PCLPLSPYDAVATASRLGPLYLLEPEGGRAYQALHCPGTGEYCKYPAQ
BALAVLEVARLPKHVARLPTEVLGATGOLIAFETRHGMSLVSRRH1PEPEAA
VLFROMATLAHQHGLVLRDLKLCRFPADRRKRLVLENLDSCLVTGPDRLMD
KHCPAYVGEPIILSSRASYSKADVMSLGVALTFTMLAGHPDPSREVLLFGKIRRG

ORIGIN
Query Match 50.6%; Score 1070.6; DB 12; Length 1077;
Best Local Similarity 99.6%; Pred. No. 3.2e-213;
Matches 1073; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
AYAPAGSAPARCLVRCILREPAERLTATGILHPMLRQDPMPLAPTRSHLWAAQ
VVPDGLIDDEAREEBEDREVLYGL"
QY 161 ATGCGAGCCACCCCTCTGCTGCTCTCTGCGGGTCCCTCTCCAGAAAGAGCGGTGGAG 220
Db 1 ATGCGAGCCACCCCTCTGCTGCTCTCTGCGGGTCCCTCTCCAGAAAGAGCGGTGGAG 60
QY 221 TTGATGACACTTGAATCCGAGCGTCCGTTCCAGAAAAGAGCTGAAGTGGGCCCAAG 280
Db 61 TTGATGACACTTGAATCCGAGCGTCCGTTCCAGAAAAGAGCTGAAGTGGGCCCAAG 120
QY 281 CCCAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 340
Db 121 CCCAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 341 GCTGTGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 400
Db 181 GCTGTGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 401 CGGCTCTACAGAGCCCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 460
Db 241 CGGCTCTACAGAGCCCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 300
QY 461 GTCCAGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 520
Db 301 GTCCAGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 521 GCTCGGCCCACTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580
Db 361 GCTCGGCCCACTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 581 CATGGGACATGACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 640
Db 421 CATGGGACATGACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 641 GTGCTCTTCCGCGAGATGGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
Db 481 GTGCTCTTCCGCGAGATGGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 701 CGTATCTCAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 760
Db 541 CGTATCTCAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 761 GAGAACTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 820
Db 601 GAGAACTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 821 GCGTGGCCAGGCTACGTGGGACCTGAGATATGAGCTACAGGGCTCTATCTCGGGCAAG 880
Db 661 GCGTGGCCAGGCTACGTGGGACCTGAGATATGAGCTACAGGGCTCTATCTCGGGCAAG 720
QY 881 GCAAGCCGATGCTGAGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 940
Db 721 GCAAGCCGATGCTGAGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 780
QY 941 TTCCAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1000
Db 781 TTCCAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 1001 CCTGCAAGGCTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1060
Db 841 CCTGCAAGGCTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 900
QY 1061 GCTGAAGGGCTCAGAGCCAGAGCATCTCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAG 1120
Db 901 GCTGAAGGGCTCAGAGCCAGAGCATCTCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAG 960
QY 1121 CCCTTAGCCCAACCCGATCCATCTCTGGAGGCTGCCAGGTGCTCTGATGAGACTG 1180

Dd		961	CCCTTAGCCCCAACCCCATCCCATCTCTGGAGGCGTCCACAGTGTCCTCGATGGACTG	1020
QY		1181	GGGCTGGACGAAGCCAGGAGAAGAGAGAGACAAGAACTGTTCGTATGGCTAG	1237
Dd		1021	GGGCTGGACGAAGCCAGGAGAAGAGAGAGAGAAGTAAGTTCTGTATAGGCTTG	1077
	RESULT 10			
	BD248318			
	LOCUS	BD248318	1085 bp	DNA linear PAT 17-Jul-2003
	DEFINITION	Human cDNA and polypeptides having kinase functions.		
	ACCESSION	BD248318		
	VERSION	BD248318.1	GI:33058088	
	KEYWORDS	JP 2002524048-A/3.		
	SOURCE	Homo sapiens (human)		
	ORGANISM	Homo sapiens		
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1085) Vitca,D.G., Bird,T.A., Anderson,D.M. and Marken,J.S. Human cDNA and polypeptides having kinase functions Patent: JP 2002524048-A 3 06-AUG-2002;		
	AUTHORS	IMRONEX CORP		
	TITLE	IMRONEX CORP		
	JOURNAL	OS Homo sapiens (human) PN JP 2002524048-A/3 PD 06-AUG-2002 PF 03-AUG-1999 JP 2000563803 PR 04-AUG-1998 US 60/095270,11-SEP-1998 US 60/099972 FI DUNE G VIRCA-TIMOTHY A BIRD DIRK M ANDERSON,JOHN S MARKEN PC C12N15/09,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N9/ PC 12,C12Q1/00, G01N33/15,G01N33/50,G01N33/566,G01N33/577//C12P21/08,C12N15/00, PC C12N5/00		
	COMMENT	CC Human cDNA and polypeptides having kinase functions FH Key		
		FT source 1..1085		
		Location/Qualifiers		
		/organism="Homo sapiens (human)". 1..1085 /organism="Homo sapiens" /mol_type="genomic DNA"		
		/db_xref="taxon:9606"		
	FEATURES			
	source			
	ORIGIN			
	Query Match	50.6%; Score 1069.8; DB 6; Length 1085;		
	Best Local Similarity	99.7%; Pred.No. 4.7e-213;		
	Matches 1082; Conservative	0; Mismatches 2; Indels 1; Gaps 1,		
QY		35	CGGGAGAGGGCTGGAGACTGGGTGGGATCCCGAGCTCCGCAGAGAGGCA-CGGGCGGGGCC	93
Dd		1	CGGGAGAGGGCTGGAGACTGGGTGGGATCCCGAGCTCCGCAGAGAGGCA-CGGGCGGGGCC	60
QY		94	ACCTGCTGATGCTCGGAGAGGCTCTAGAGCCCGAGCGGCGCCCGGGCCCAAGCGGAACGACG	153
Dd		61	ACCTGCTGATGCTCGGAGAGGCTCTAGAGCCCGGCGGCGCCCGGGCCCAAGCGGAACGACG	120
QY		154	GGGCGAGATGCGAGCCACCCTCTGGCTGCTCTCGGAGGTTCCCTGTCCAGGAAGAAGCG	213
Dd		121	GGGCGAGATGCGAGCCACCCTCTGGCTGCTCTCGGAGGTTCCCTGTCCAGGAAGAAGCG	180
QY		214	GTTGAGTTGATGATGACAACTTAGATACCGAGAGGTCGCCGTCAGAAACGAGCTCGAAGTGG	273
Dd		181	GTTGAGTTGATGATGACAACTTAGATACCGAGAGGTCGCCGTCAGAAACGAGCTCGAAGTGG	240
QY		274	GCCCAGCCCAAGCTGCCCCCTGCTGTTGCCCTGAGCCCAACTACTGCTCCAGANTCG	333
Dd		241	GCCCAGCCCAAGCTGCCCCCTGCTGTTGCCCTGAGCCCAACTACTGCTCCAGANTCG	300
QY		334	TGCAACTGCTGTGGGCACTGCTCCCGTCTTGGGCGCCATTATGTCCTCTGGAGCCCGAGGA	393
Dd		301	TGCAACTGCTGTGGGCACTGCTCCCGTCTTGGGCGCCATTATGTCCTCTGGAGCCCGAGGA	360

QY	394	GGGCGGGGGGGCCCTACCGAGGCGCTGCACTGGCCCTTACAAGGCACTGAGATATCCGCAAGGT	453
Db	361	GGGCGGGGGGGCCCTACCGAGGCGCTGCACTGGCCCTTACAAGGCACTGAGATATCCGCAAGGT	420
QY	454	GTAACCCGATCCAGAGAGCCCTGGCCGCTGAGAGCCCTTACGCGCGGCTGCCCCGACAA	513
Db	421	GTAACCCGATCCAGAGAGCCCTGGCCGCTGAGAGCCCTTACGCGCGGCTGCCCCGACAA	480
QY	514	GCATGTGGCTGGGCGCCATGAGGTCCCTGGGTGTGTATCCAGACTCTCTAGCGCTTTTTCAC	573
Db	481	GCATGTGGCTGGGCGCCATGAGGTCCCTGGGTGTGTATCCAGACTCTCTAGCGCTTTTTCAC	540
QY	574	TCCGAGCCCATGGGGAGACATGACAGCCCTGATGCGAAGCGCACCGATATCCCTGAGCCCTGA	633
Db	541	TCCGAGCCCATGGGGAGACATGACAGCCCTGATGCGAAGCGCACCGATATCCCTGAGCCCTGA	600
QY	634	GGCTGGCGTCTCTTCCGCGAGATGCGACCGCCCTTGGCGCATGTGTATCCAGACAGGTC	693
Db	601	GGCTGGCGTCTCTTCCGCGAGATGCGACCGCCCTTGGCGCATGTGTATCCAGACAGGTC	660
QY	694	GGTCCCTGGGATCTCAAGCTGTGTGGCTTTGTCTTGCTGACCGGTGAGAGAAAGACT	753
Db	661	GGTCCCTGGGATCTCAAGCTGTGTGGCTTTGTCTTGCTGACCGGTGAGAGAAAGACT	720
QY	754	GGTCTGGAGAAACCTGAGAGACTCTTGCGTGTGACTGAGGCGCAGATATTTCCCTGTGGGA	813
Db	721	GGTCTGGAGAAACCTGAGAGACTCTTGCGTGTGACTGAGGCGCAGATATTTCCCTGTGGGA	780
QY	814	CAGACAGCGCTGCCAGACTTCTGTGGGACCTGAGATATCTAGCTTACAGGGGCTTATATCTC	873
Db	781	CAGACAGCGCTGCCAGACTTCTGTGGGACCTGAGATATCTAGCTTACAGGGGCTTATATCTC	840
QY	874	GGGCAAGGCAAGCGGATGTCTGGAGACTTGGGGGCTGGGCGCTCTTCAACATGTGCGCGGCCA	933
Db	841	GGGCAAGGCAAGCGGATGTCTGGAGACTTGGGGGCTGGGCGCTCTTCAACATGTGCGCGGCCA	900
QY	934	CTACCCCTTCAGAGACTGAGGCTGTCTGTCTTTCGGCAAGATCCGCGGGGGCCCTA	993
Db	901	CTACCCCTTCAGAGACTGAGGCTGTCTGTCTTTCGGCAAGATCCGCGGGGGCCCTA	960
QY	994	CGCCTTGCTGCAAGGCTCTTGGGCGCCCTGCGCCGCTGTCTGGTTGGCTGGGCTCTTGGTCG	1053
Db	961	CGCCTTGCTGCAAGGCTCTTGGGCGCCCTGCGCCGCTGTCTGGTTGGCTGGGCTCTTGGTCG	1020
QY	1054	GGAGCCAGCTGAACGGCTCAACAGGCAAGGCACTCTCTGCAACCCCTGAGCTGCGACAGGA	1113
Db	1021	GGAGCCAGCTGAACGGCTCAACAGGCAAGGCACTCTCTGCAACCCCTGAGCTGCGACAGGA	1080
QY	1114	CCCGGA 1118	
Db	1081	CCCGGA 1085	
RESULT 11			
AR411700			
LOCUS	AR411700	1074 bp	DNA
DEFINITION	Sequence 9 from patent US 6638721.	linear	PAT 18-DEC-2003
ACCESSION	AR411700		
VERSION	AR411700.1	GI:40164081	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1074)		
AUTHORS	Meyers, R., Kapeller-Libermann, R. and Williamson, M.		
TITLE	Human protein kinases and uses therefor		
JOURNAL	Patent: US 6638721-A 9/28/OCT-2003;		
FEATURES	Location/Qualifiers		
source	1..1074		
	/organism="unknown"		
	/mol_type="genomic DNA"		

FEATURES	source
LOCUS	AF250311
DEFINITION	Homo sapiens SKIP3 mRNA, complete cds.
ACCESSION	AF250311
VERSION	AF250311.1 GI:14276268
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1083)
TITLE	Kiss-Toth E., Wyllie D.H., Qvarnstrom E.E. and Dower S.K. Identification of pro-inflammatory cytokine signalling network components by transcription expression screening
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1083)
AUTHORS	Kiss-Toth E., Wyllie D.H., Qvarnstrom E.E. and Dower S.K.
TITLE	Direct Submission
JOURNAL	Submitted (29-MAR-2000) Division of Molecular & Genetic Medicine, University of Sheffield, Royal Hallamshire Hospital, Floor M, Sheffield S10-2UF, UK
FEATURES	Location/Qualifiers
source	1..1083
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	1..1083
	/codon_start=1
	/product="SKIP3"
	/protein_id="AAK58175.1"
	/db_xref="gi:14276268"
	/translation="MRATPIIAPAGLSLRKKRIELDNDJTERPYOKRARGSDPDRLEPCILDSPRTADRTAVATASRLGPIYLLEPBBGARGAYKALHCGTREVYCKYPCVLEALVEPARVPKHKVAPRVEVLGQULFAFTRTHGDHSLVSRHRIIPERPAVLFRVMAVLAHACHQHGLVLRDLKICRFVPRDRERKKLVLENLEDSCVLTPGDSIS

ORIGIN

WDKHAACEPAYVGEPEILSSRASYSYGKADVMSLGVALLFTMLAGHPYEQDSEPVLLFGKIR
RGAYALPAGLSAPARCLVRCLLRREPAERTLTATGILHPWLQDMPPLAPTRSHLWEA
AQVVPDGLGLDEAREEEGDREVLVYG"

Query Match	49.8%	Score 1053	DB 9	Length 1083
Best Local Similarity	99.0%	Pred. No. 1.5e-209		
Matches 1072; Conservative	0	Mismatches 5	Indels 6	Gaps 1

QY	161	ATGGAGCCACCCCTCGAGGTGCTCCGCGGGTTCCTGTGCAGAAAGAGCGGTGAG	220
Db	1	ATGGAGCCACCCCTCGAGGTGCTCCGCGGGTTCCTGTGCAGAAAGAGCGGTGAG	60
QY	221	TTGGATGACAACTTAGATACCGAGCGTCCGCTCCAGAAAGAGGTGCAATGGGCCCCAG	280
Db	61	TTGGATGACAACTTAGATACCGAGCGTCCGCTCCAGAAAGAGGTGCAATGGGCCCCAG	120
QY	281	CCCGAGCTGGCCCCCTGCTGTGTCGCTCCCTGAGGCCACTACTGTCTCCAGATCTGTCAACT	340
Db	121	CCCGAGCTGGCCCCCTGCTGTGTCGCTCCCTGAGGCCACTACTGTCTCCAGATCTGTCAACT	180
QY	341	GCTGTGGCCACTGCTCCCTGTGGGCCCTATGTCTCTGTGAGGCCGAGGAGGCGGG	400
Db	181	GCTGTGGCCACTGCTCCCTGTGGGCCCTATGTCTCTGTGAGGCCGAGGAGGCGGG	240
QY	401	CGGCGCTTACCAAGGCTCTGCACTGCTCCCTTACAGGCACTAGATTTACTCTCAAGTGTACCC	460
Db	241	CGGCGCTTACCAAGGCTCTGCACTGCTCCCTTACAGGCACTAGATTTACTCTCAAGTGTACCC	300
QY	461	GTCAGAGAAAGCCCTGGCGGTGTGTGGAGCCCTACGCGGGCTGGCCCCGCAACAATGTG	520
Db	301	GTCAGAGAAAGCCCTGGCGGTGTGTGGAGCCCTATGCGGGGTGCCCCCGCACAGCAATGTG	360
QY	521	GCTCGGCGCACTGAGGTCTGTGGCTGTATACCCAGCTCTCTAAGCTTTTTCATCTGAGCC	580
Db	361	GCTCGGCGCACTGAGGTCTGTGGCTGTATACCCAGCTCTCTAAGCTTTTTCATCTGAGCC	420
QY	581	CATGGGGACATGACACAGCTGTGTGTGCAACGCGCACCGATTCCTTGAGCTTGAGGCTGCC	640
Db	421	CATGGGGACATGACACAGCTGTGTGTGCAAGCGCGCACCGATTCCTTGAGGCTTGAGGCTGCC	480
QY	641	GTGCTTTCCGCGCAGATATGGCCACCGCCCTGGCGCACTGTACACAGCACCGGTGTGTCTG	700
Db	481	GTGCTTTCCGCGCAGATATGGCCACCGCCCTGGCGCACTGTACACAGCACCGGTGTGTCTG	540
QY	701	CGTGAATCAAGCTGTGTGTGTCTTTGTCTTGC-----TGAACGATGAGGAGAAAGCTG	754
Db	541	CGTGAATCAAGCTGTGTGTGTGTCTTTGTCTTGCCTGACCGTGAACCGTGAAGAAAGAGCTG	600
QY	755	GTGCTGAGAACTGTGAGAGACTCTCTGCGTGTGACTGTGGGCAAGATGATTCCTGTGGGAC	814
Db	601	GTGCTGAGAACTGTGAGAGACTCTCTGCGTGTGACTGTGGGCAAGATGATTCCTGTGGGAC	660
QY	815	AAGCACGGTGTGCCCACTAGTGGGAACTTAGATATCTACAGTCACGGAGCTCATATCTCG	874
Db	661	AAGCACGGTGTGCCCACTAGTGGGAACTTAGATATCTACAGTCACGGAGCTCATATCTCG	720
QY	875	GCGAAGGAGCGAGTGTGTGAGGCTGTGGCGGTGTGGGCTTTCAACCATGTGTGCGCGGAC	934
Db	721	GCGAAGGAGCGAGTGTGTGAGGCTGTGGCGGTGTGGGCGGTCTTCAACATGTGTGCGCGGAC	780
QY	935	TACCCCTTCCAGAGTCTGGAGCTGTGTCTGTCTTTGTGGCAAGATTCGCGGGGGGCTTAC	994
Db	781	TACCCCTTCCAGAGTCTGGAGCTGTGTCTGTCTTTGTGGCAAGATTCGCGGGGGGCTTAC	840
QY	995	GCGTTGCTGAGAGGCTGTGGGCCCTTGGCCCGGTGTGATTTGGTGTGCTCTTGTGTGG	1054
Db	841	GCGTTGCTGAGAGGCTGTGGGCCCTTGGCCCGGTGTGATTTGGTGTGCTCTTGTGTGG	900
QY	1055	GAGCAGCTGAAAGGCTCTACAGGCCACAGGACATCTCTTGCAACCCCTGTGCTGTGACAGAC	1114
Db	901	GAGCAGCTGAAAGGCTCTACAGGCCACAGGACATCTCTCTGCAACCCCTGTGCTGTGACAGAC	960


```
Db      196 CTAGAGCCCGGCGAGGCGTGGAGCTGGGCTGGGATCCGAGCTCGGCAGAGCGCA -GNG 254
QY      87  CCGGCCCACTGCTGTGTGTCCTGAGGCTCTGAGCCCGGCGGGCCCGGCCCACTCGG 146
Db      255 CCGGCCCACTGCTGTGTGTCCTGAGGCTCTGAGCCCGGCGGGCCCGGCCCACTCGG 314
QY      147 AACGACGGGGCGAGATGGAGCAACCCCTGTGGCTGCTCTGTGGGTTCCCTGTCCAGGA 206
Db      315 AACGACGGGGCGAGATGGAGCAACCCCTGTGGCTGCTCTGTGGGTTCCCTGTCCAGGA 374
QY      207 AGAAGCGGTTGGAGTTGATGACAACTTAGATACGAGCGTCCCGTCCAGAAACGAGCTC 266
Db      375 AGAAGCGGTTGGAGTTGATGACAACTTAGATACGAGCGTCCCGTCCAGAAACGAGCTC 434
QY      267 GAAGTGGGCCCGAGCCAGACTGCCCCCTGTGCTGTGGCCCTGAGCCCACTACTGCTC 326
Db      435 GAAGTGGGCCCGAGCCAGACTGCCCCCTGTGCTGTGGCCCTGAGCCCACTACTGCTC 494
QY      327 CAGATCTGTGCACTGTGTGGGCACTGGCTCCCGTCTGGGCGCTATGTCTCTGAGAGC 386
Db      495 CAGATCTGTGCACTGTGTGGGCACTGGCTCCCGTCTGGGCGCTATGTCTCTGAGAGC 554
QY      387 CCGAGAGGGCGGGCGGGGCTTACAGAGCCCTGCACTGCCCCCTACAGGCACTGATATACCT 446
Db      555 CCGAGAGGGCGGGCGGGGCTTACAGAGCCCTGCACTGCCCCCTACAGGCACTGATATACCT 614
QY      447 GCAAGGTGTACCCCGTCCAGGAAGCCCTGCGCTGTGAGCCCTACGCGCGGCTGCCCC 506
Db      615 GCAAGGTGTACCCCGTCCAGGAAGCCCTGCGCTGTGAGCCCTACGCGCGGCTGCCCC 674
QY      507 CGCACAAGCATGTGTGCTGGGCCCACTGAGGTCTGTGTTACCAAGCTCTCTACGCCT 566
Db      675 CGCACAAGCATGTGTGCTGGGCCCACTGAGGTCTGTGTTACCAAGCTCTCTACGCCT 734
QY      567 TTTTCACTCGGACCCATGGGGACATGACAGCCTGTGTGGAAACGCGCAACCGTATCCCTG 626
Db      735 TTTTCACTCGGACCCATGGGGACATGACAGCCTGTGTGGAAACGCGCAACCGTATCCCTG 794
QY      627 AGCTGAGGCTGCGGTCTTTCGCGAATGGGCCACCGCCTGCGCACTGTACACAGC 686
Db      795 AGCTGAGGCTGCGGTCTTTCGCGAATGGGCCACCGCCTGCGCACTGTACACAGC 854
QY      687 ACGGTCTGTGCTGCGGATCTCAAGCTGTGTGCTTTGTCTTGCTGACCGTGAAGGA 746
Db      855 ACGGTCTGTGCTGCGGATCTCAAGCTGTGTGCTTTGTCTTGCTGACCGTGAAGG- 913
QY      747 AGAAGCTGTGTGAGAACTGAGAGACTCTGCGTGTGACTGAGC 794
Db      914 AGAAGCTGTGTGAGGA--CTGAGAGACTGTGCGTGTGACTGAGC 959
```

Search completed: August 29, 2004, 10:31:30
Job time : 8351 secs